AL/CF-SR-1995-0018

ARMSTRONG LABORATORY

INVESTIGATION OF EYE AND HEAD CONTROLLED CURSOR POSITIONING TECHNIQUES (U)

Joshua Borah

APPLIED SCIENCE LABORATORIES
175 MIDDLESEX TURNPIKE
BEDFORD MA 01730

19960206 035

SEPTEMBER 1995

FINAL REPORT FOR THE PERIOD 17 MAY 1993 TO 30 JUNE 1995

Approved for public release; distribution is unlimited

AIR FORCE MATERIEL COMMAND
WRIGHT-PATTERSON AIR FORCE BASE, OHIO 45433-6573

NOTICES

DISCLAIMER

This Special Report is published as received and has not been edited by the Technical Editing Staff of the Armstrong Laboratory.

TECHNICAL REVIEW AND APPROVAL

AL/CF-SR-1995-0018

This report has been reviewed by the Office of Public Affairs (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.

The voluntary informed consent of the subjects used in this research was obtained as required by Air Force Instruction 40-402.

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER

Kenneth R. Boff

KENNETH R. BOFF, Chief
Human Engineering Division
Armstrong Laboratory

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
September 1995

3. REPORT TYPE AND DATES COVERED
Final Report, 17 May 93 to 30 June 95

4. TITLE AND SUBTITLE
Investigation of Eye and Head Controlled
Cursor Positioning Techniques (U)

5. FUNDING NUMBERS
C F41624-93-C-6006
PE 62202F
PR 3005
TA CH
WU 3B

6. AUTHOR(S)
Joshua Borah

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Applied Science Laboratories
175 Middlesex Turnpike
Bedford MA 01730

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Armstrong Laboratory, Crew Systems Directorate
Human Systems Center
Air Force Materiel Command
Wright-Patterson AFB OH 45433-7022

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
AL/CF-TR-1995-0018

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*
The project objective was to investigate the feasibility of using eye point-of-gaze and head control of a display cursor in place of, or to supplement, manual control for cursor positioning tasks. Of particular concern was the problem of positioning a cursor with respect to targets that may need to be small with respect to eye tracker accuracy and precision. The approach to this problem was to incorporate a second control modality, other than eye movement, for closed loop error correction (i.e. fine tuning). A second objective was to investigate the applicability of Fitts' law and related movement time prediction models to eye control.

Several techniques were subjectively tested which allow users to "close the loop" and fine tune point-of-gaze controlled cursor positions. All but one of these techniques were rejected on the basis of preliminary subjective tests. A technique that allowed the user to switch from point-of-gaze control to a low gain head position control when near the target seemed very effective during informal, subjective tests. This technique was used in a set of formal experiments, along with pure point-of-gaze control, pure head position control, and standard mouse control of display cursor position. The latter two control techniques have been studied fairly extensively by others, and data using these techniques help relate current results to previous work.

All four of the above control techniques were formally tested using two serial, cursor positioning tasks. One of the tasks included a search component, and was designed to represent a realistic computer interface task. The other did not have a search component, and was designed to facilitate analysis of motion time with Fitts' law and related models.

14. SUBJECT TERMS
Eye, Eye Movements, Control,
Tracking, Head

15. NUMBER OF PAGES
48

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UNLIMITED

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

THIS PAGE INTENTIONALLY LEFT BLANK

INTRODUCTION AND SUMMARY

The overall project objective was to investigate the feasibility of using eye point of gaze, and head control of a display cursor, in place of, or to supplement, manual control for cursor positioning tasks. Of particular concern, was the problem of position a cursor with respect to targets that may need to be small with respect to eye tracker accuracy and precision. The approach to this problem was to incorporate a second control modality, other than eye movement, for closed loop error correction (i.e. fine tuning). A second objective was to investigate the applicability to eye control, of Fitts' law, and related movement time prediction models.

Several techniques were subjectively tested, for allowing users to "close the loop", and fine tune point of gaze controlled cursor positions. All but one of these techniques were rejected on the basis of preliminary subjective tests. A technique that allowed the user to switch from point of gaze control to a low gain head position control, when near the target, seemed very effective during informal, subjective tests. This technique was used in a set of formal experiments, along with pure point of gaze control, pure head position control, and standard mouse control of display cursor position. The latter two control techniques have been studied fairly extensively by others, and data using these techniques help in relating current results to previous work.

All four of the above control techniques were formally tested using two, serial, cursor positioning tasks. One of the tasks included a search component, and was designed to represent a realistic computer interface task. The other did not have a search component, and was designed to facilitate analysis of motion time with Fitts' law and related models.

Mouse control produced the fasted median task completion times, if we consider the entire range of target sizes and distances. Pure eye control had significantly faster median times when the accuracy and precision limitations of the eye tracker were not exceeded, but these limits restrict the technique to relatively large target sizes. When targets are smaller than eye tracker performance limits, pure eye control becomes impossible.

Both pure head control, and eye control with head controlled fine tuning, proved viable. It is possible to reliably select targets across the entire range of target sizes and motion distances tested. Both techniques produced significantly slower median motion times than mouse control, or pure eye control with large targets. Contrary to expectations from subjective tests, pure head control yielded median motion times that were equivalent to, or slightly faster than the technique employing eye control with head controlled fine tuning ("eye&hd"). The eye&hd technique did produce minimum (as opposed to median) times that were similar to those for mouse control, but the eye&hd data show much greater variance than mouse or pure head control data.

Although all control techniques show a significant Fitts' law relationship, variance increases with index of difficulty (heteroscedasticity). A better fit is obtained by taking log of task completion time as the predicted variable, rather than task completion time. Variance characteristics of the data may be

influenced by the serial nature of the task, and by transport delays in the control mechanism. Fitts' law parameters from several previous studies of mouse, head, and eye movement control are presented for comparison.

The pog&hd control paradigm is clearly viable, but based just on task completion time measures, does not show an advantage over pure head control. Methods are discussed for more thorough analysis of the existing data, and for possible improvements in the pure eye, and pog&hd control techniques.

BACKGROUND

Eye Movement Measurement Equipment

There are several different eye tracking techniques currently in use for humans. When requirements call for relatively unobtrusive point of gaze measurement, the practical possibilities are quickly reduced to optical techniques that measure the relative position of two features on the eye ball. The two features generally used are the pupil, and the reflection of a light source on the outer surface of the cornea (the "corneal reflection", or CR). Furthermore, if the environment requires a lot of head motion, or if measurements must be made over a very wide field of view, the current state of the art will probably dictate use of a head mounted device. More thorough treatments of eye tracking techniques can be found in Young and Sheena (1988) and Borah (1989).

Currently available systems, using the pupil to CR technique, work as follows. A near infra red light source illuminates the eye, and the eye is viewed by solid state video sensor. The camera video signal is treated by some form of image processor to find the center of the pupil and CR. This is essentially a pattern recognition task. The processor must then use the pupil center to CR vector to compute eye line of gaze with respect to the optics.

Pupil to CR type systems measure eye position with respect to the eye tracker optics. If the eye tracker optics are head mounted, and the desired output is point of gaze with respect to the room, or cockpit, then head position and orientation must also be measured.

The most accurate and versatile, commercially available, head tracking devices use a magnetic technique. The position and orientation of a small magnetic sensor is determined, in all 6 degrees of freedom, with respect to a magnetic transmitter. When used as a head tracker, the sensor is fastened to the user's headgear, and the transmitter is fastened to the room, or cockpit. Being magnetic devices, they are adversely affected by metal in the environment, and by electro-magnetic emission. They can be compensated for these affects, and such devices have been successfully used in military aircraft. Optical, and ultrasonic head tracking devices also exist, but are either not commercially available, or do not yet have the required accuracy.

When eye tracker optics are head mounted, the system processor accepts input from both eye tracking, and head tracking subsystems, to determine point of gaze with respect to the environment.

When eye tracker optics are fixed to the room, or cockpit, independent head tracking is not required for the point of gaze measurement. Head motion can be accommodated by reflecting the optical path from a moving mirror. or by mounting the entire optics package on a moving platform. When the system detects that the pupil image is off center with respect to the camera field, the mirror, or moving platform, is commanded to move in the appropriate direction to re-center the image. A wide angle camera, or separate head tracking system are sometimes used to facilitate the mirror, or platform tracking task. Floor mounted (as opposed to head mounted) systems are less obtrusive, but are more constrained by optical path obstructions, head motion limits, and field of view limits; and can also be prone to data losses due to failure of the moving mirror or optics package to keep up with the head.

Pupil to CR systems are usually capable of measuring eye line of gaze to an accuracy of about 1 degree visual angle. Jitter of the measurement, during an eye fixation, is usually on the order of 1/4 to 1/2 degree visual angle, and resolution (the smallest eye rotation that can be sensed) also tends to be 1/4 to 1/2 degree visual angle. For such systems, update rate is generally 60 Hz (if using American standard video), and there is a data transport delay of about 50 msec (3 video fields).

Non standard video sensors, that update at faster than 60 Hz rates, are available, and can be used to increase the update rate of pupil to CR type systems. The higher update rate usually exacts a performance price in the form of reduced system robustness (ability to find the critical features in very dim or cluttered images), and reduced spatial resolution.

Eye Movement as a Computer Interface

The keyboard, mouse, and trackball are currently among the most common human-computer interface devices for user input. All must be operated manually. In certain environments, such as aircraft cockpits, the operator's hands are very busy with other tasks, and it would be beneficial to have alternate means for computer interface. Even when the hands are not quite so overloaded, more efficient and more natural interface techniques would be of obvious benefit. Techniques that incorporate eye movement measurement are obvious candidates for consideration.

Use of eye position as a control input device presents some unique problems. The accuracy of currently available eye tracking devices, which are suitably unobtrusive, is about 1 degree visual angle. Visual acuity is far better than this, and size and spacing of elements on a computer display usually involves separations significantly smaller than 1 degree visual angle. Even with an instrument that could measure eye ball position with infinite accuracy, there is a physiological limit to how well we can know what a person is looking at by knowing which way the eye ball is pointing. Although the precise value of this limit is not clear from available literature, it could easily be on the order of 1/8 degree visual. Within such a limit, visual attention can be shifted to different areas inside the fovea (high acuity area of the retina) without corresponding eye ball motion.

Exerting control over something by consciously directing gaze is not a usual or natural behavior. In fact, scan pattern is usually determined subconsciously. If an eye control task is not designed properly, it may feel very unnatural and annoying.

Jacob (1990) considers eye tracking as it relates to the overall issue of human/computer interaction. His philosophical approach is to design interaction protocols that make use of natural human behaviors, rather than requiring learned techniques. He incorporated eye control in an system designed to allow users to position objects on a display (e.g. icons representing ships on a map) and to call up information about individual objects. A fixation filter and an accuracy enhancement technique was used to post process data from the eye tracker. His experiment was primarily subjective, and qualitative. He found that when the eye tracker interface was working well it seemed as though the system were "reading the user's mind". On the other hand, when the eye tracker performance was not precise enough, it could be terribly frustrating.

In the work presented in this report, we assume that it may sometimes be acceptable to require a learned behavior, especially if necessary to meet another requirement, such as freeing up the user's hands. With this in mind, we have attempted to enable control tasks that may exceed the open loop accuracy of an eye position measurement system. Furthermore, we have not attempted to design a human computer interaction system; rather we have restricted ourselves to investigating performance of a simple cursor positioning and target selection task. The word "cursor" is used, herein, to refer to a computer program's knowledge of the user's point of interest on a display. This point may or may not be represented by a visible cursor.

Fitts' Law

Fitts (1954) extrapolated from an information transmission theory model (Shannon and Weaver, 1949) to propose a model for predicting human movement times. The usual form of Fitts' law is

$$MT = a + b \cdot ID \qquad (1)$$

where *MT* is motion time, *ID* is index of difficulty, and *a* and *b* are constants. The index of difficulty proposed by Fitts is

$$ID = \log_2 (2D / W) \qquad (2)$$

where D is distance to the target and W is target width. The choice of base 2, for the logarithm, is arbitrary, but allows index of difficulty units to be expressed as "bits". Fitts law, and variations of Fitts law have since proven to be useful models for a wide variety of tasks. Even in situations where Fitts's law does not provide the best possible modeling fit, it is often a useful means for comparison across different types of movement tasks.

It has often been noted that Fitts law, as stated in equations 1 and 2, may have problems at very low ID values. Specifically, data often show larger than

predicted motion times at very low values of ID. To correct this tendency, Welford (1960) proposed a revised index of difficulty.

$$ID_{Welford} = \log_2(D/W + 0.5) \tag{3}$$

MacKenzie (1992) has proposed an index of difficulty formulation based on a more rigorous analogy to Shannon's information transmission theorem.

$$ID_{Shannon} = \log_2(D/W + 1.0) \tag{4}$$

The Welford and Shannon index of difficulty formulations, have often been found to produce better correlation's, when used in equation 1, than the ID first used by Fitts.

Sherridan and Ferrell (1963) tested a remote controller device with a transmission delay. In this case, the logarithm of motion time, rather than time itself, showed the best linear relation to index of difficulty.

Of particular relevance to the current study, are previous studies of Fitts law as applied to mouse, head movement, and eye movement control. Epps (1986); Card, English, and Burr (1978); Radwin (1990); and Lin, Radwin, and Vanderheiden (1992) all found Fitts law to be an adequate predictor of motion time for mouse control. All, except Epps, used clearly discrete tasks. In the Card, et. al., Radwin, and Lin, et. al. studies, subjects were given time to prepare for each trial and started from a set cursor position. In the Epps study, each trial started from the center of the preceding target, and each trial was triggered by successful completion of the preceding trial.

Card, et. al. required subjects to move a cursor to a highlighted word, in a block of text, and to confirm the selection with a button press. Epps required movement of a cross hair to within a target square, followed by a confirming button press. Radwin, and Lin, et. al. tested cursor movement from a central home position to circular targets, and used on-target dwell time of 62.5 ms, rather than a button press, to confirm target selection.

Epps, Radwin and Lin et. al. all fit data to Fitts ID, while Card, et. al. used the Welford formulation.

The Lin, et. al. study was designed to explore the affects of device gain, and found that optimal mouse gain (mouse_movement / cursor_movement) was about 2.0. The mouse acceleration feature was disabled so that control was a purely linear function.

Radwin (1990), and Lin, et. al. (1992), also tested a head controller with the same experimental paradigm used for the mouse. The head controller was

based on an ultrasonic device. Lin, et. al. varied gain for the head as well as mouse controller, and found optimal head control gains of between 0.3 and 0.6. Jagacinski and Monk (1985) tested a helmet mounted sight, and found adequate motion time predictions using the Welford ID formualtion. The helmet sight, which used an optical technique to detect helmet position, was used to control a display cursor. The discrete task required movement of the cursor from a central home position to a circular target, and confirmation by on-target dwell time (344 ms).

Ware and Mikaelian (1987) used a floor mounted eye tracker, employing the pupil to CR technique described in the previous section, to control a display cross hair. The discrete task required subjects to begin each trial by fixating a central target, and to move the cursor to a highlighted rectangle. Several confirmation techniques were used, including on-target dwell time (400 msec), and a button press. Regressions were plotted of confirmation time versus the Welford ID. The range of ID values was very small (-1 to 1.8 bits), however, because of the requirement to keep targets larger than system accuracy. Note that the negative ID results from a case in which the cursor starts from the center of the target to be selected (zero motion distance). There must still, of course, be a finite time before the target is confirmed.

There is some inconsistency among studies in the way data is used for regression analyses. The original Fitts taping task Fitts (1954), for example, was analyzed by taking the mean, for all subjects, at each different target size and motion distance combination, and performing a regression on the mean data. Epps (1986) based regressions on mean motion times for each individual subject at each target size and distance combination. Radwin (1990) based regressions on mean data for each ID value, pooled across subjects. Individual regressions were computed, however, for each different radial motion direction. Jagacinski and Monk (1985) used median values at each condition. Ware and Mikaelian (1987) plotted regressions based on all data points. In some papers it is not clear exactly what data points were used for regression. Use of raw data, versus mean or median data, will often have little affect on the regression coefficients, but may have an enormous effect on the correlation coefficient.

The studies cited above are by no means an exhaustive list of work relating to Fitts law. For a thorough review of Fitts' law theory and research, especially as relates to human computer interaction, see MacKenzie (1992).

Fitts' law is not the best model for saccadic eye movements. Saccades are the rapid eye ball rotations that move gaze from one fixation point to another, and during which little if any visual information is acquired. A very clear linear relationship has been shown relating saccadic duration to saccade distance, at least for saccades of less than 15 degrees visual angle. Abrams, Meyer, and Kornblum (1989), for example, show a linear relation ($r = 0.99$) between mean saccade duration and mean saccade amplitude, for saccades of between 3 and 9 degrees visual angle. Their data fit the model

$$ST = k_0 + k_1 \cdot SA \tag{5}$$

where *ST* is saccade duration, *SA* is saccade amplitude, k_0 is 23.6 msec, and k_1 is 2.94 msec/degree. Note that this is not a model for performance of a control task relying on eye position measurement, rather it is strictly a relation between duration and distance for a saccadic eye movement.

METHOD

Equipment

An experimental set up was arranged with an ASL series 4000, head mounted, eye and head tracker as the central component. The system includes head mounted optics, a magnetic head tracker, an eye tracker computer (80486 PC), EYEHEAD integration software, and a stationary scene camera.

The equipment configuration is shown, schematically, in figure 1.

Head band mounted optics illuminate the eye with a near infra red beam, and image the eye onto a video sensor, by reflecting from a visor. The visor is coated to be reflective in the near infra-red, but transmissive in the visible spectrum, so that it appears clear to the subject. The video image is processed to identify the pupil, and the light source reflection from the cornea (CR), and the pupil to CR technique (as described in the Background section) is used to

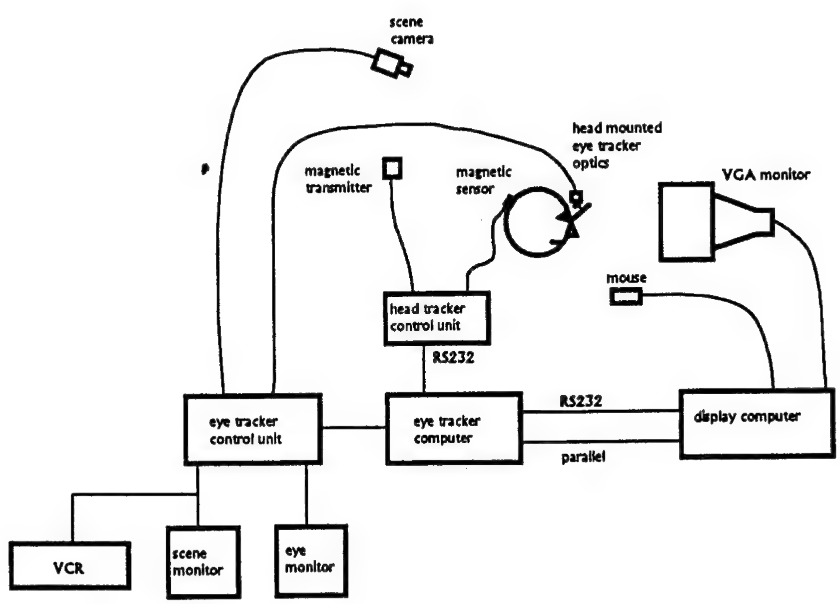


Figure 1. Equipment configuration

compute line of gaze with respect to the head gear.

The system is equipped with a magnetic position sensing system (Polhemus Navigation Science, 3 Space Tracker). The small magnetic sensor is fastened to the subject's head band, and, for this study, the magnetic transmitter was fastened to a post, behind the subject's chair. EYEHEAD integration software, running on the eye tacker computer, determines the position of the line of gaze vector, in space; determines which of several pre-defined, room fixed surfaces are intersected by line of gaze; and determines the intersection point on that surface. The point of gaze information can be recorded on the eye tracker computer, at a rate of 60 updates per second, and sent through a serial port to another device with the same update rate. It can also be displayed, as cursor or cross hair, superimposed on a video image showing the pre-defined, room fixed, surfaces.

The eye-head tracker system must be calibrated, for each subject, by having the subject look, sequentially, at nine target points, with known positions on one of the pre-defined surfaces. Subjects must hold their heads still during the calibration (about 15 seconds). Best calibration results are achieved when the calibration target pattern subtends between 30 and 40 degree visual angle.

The system includes a test mode for which live eye tracker information is not used, rather the eye is assumed to maintain a fixed straight and level position with respect to the head. This mode was used for head control tasks.

For the current study, a 13 inch VGA monitor was positioned in front of the subject's chair, so that it was approximately 28 inches from the subject's eyes. The monitor was driven by a second 80486 PC, which will be referred to as the display computer. A mouse, also connected to the display computer, was positioned for comfortable use when the subject used the chair arm rest.

The display computer was connected to the eye tracker computer via both a serial and parallel port. The serial connection was between COM ports on both computers. The parallel connection was between the display computer printer port, and a 24 bit parallel port (ASL system "XDAT" feature) on the eye tracker computer.

The display monitor bezel was equipped with Velcro tabs, and alignment marks, so that a flat Plexiglas plate could be quickly fastened to the front of the display in a reproducible position. The Plexiglas plate extended several inches beyond the edges of the monitor bezel, and was marked with a grid pattern, and with a square pattern of 9 black circles used as eye tracker calibration points. At a 28 inch eye to screen distance, the calibration pattern subtended about 30 degrees visual angle.

The surface defined to the EYEHEAD integration system was the Plexiglas screen. The Plexiglas screen was always removed after subject calibration, but the point of gaze reported by the eye-head tracker system was the computed intersection point on the Plexiglas screen. Note that this was almost, but not precisely, coincident with the slightly curved monitor surface.

A video camera was mounted to the same post as the magnetic transmitter, so that it had a view of the display monitor from above and behind the subject's

head. A point of gaze cursor was superimposed on this image by the eye-head tracker system, and the resulting video signal was connected to a VCR.

The experiment tasks were programmed on the display computer, and task performance was timed by and recorded on this computer. The display computer requested and received data from the eye-head tracker, through the serial interface, at a rate of 60 updates per second. Flag values indicating trial number, trial start point, and trial completion point, were sent from the display computer printer port, to the parallel external data port (XDAT) that is part of the eye tracker system. These flags appear on point of gaze data recorded separately by the eye tracker computer, and allow data files from the two computers to be time synchronized.

When point of gaze information was required, the display computer program used a separate offset and gain factor, for each axis, to convert eye-head tracker point of gaze information to VGA monitor pixel coordinates. For some of the cursor control algorithms tested, the information used was actually eye position with respect to the head, rather than point of gaze on the display surface. Both types of information are available from the eye-head tracker system.

The eye tracker has a transport delay of 50 msec. The head tracker delay is not as easy to define because dynamic filters introduce varying time constants depending on angular accelerations. For the relatively small and slow head motions of seated subjects, in the experiments reported herein, it is likely that head tracker lag was comparable to that of the eye tracker. An additional delay of up to 17 msec is required for receipt of data and display update, by the display computer. It is therefore estimated that delay between point of gaze events and corresponding changes on the task display screen were about 67 msec (4 video fields). In the case of mouse control, the delay between mouse motion and corresponding display change was no more than 17 msec. Mouse button presses were detected, and acted upon, by the display computer, within 17 msec, for all types of control task.

Tasks

Two types of task were used. One, referred to as the "search task" was designed to simulate a common type of computer interaction activity, requiring control of a cursor. The other, referred to as "Fitts task" was designed to facilitate Fitts' law type analysis of performance data.

Fitts task

A trial set was initiated by a left mouse button press, causing a white circle to appear at one of 20 possible positions on the display screen. The possible positions were arranged in an even grid of 5 columns and 4 rows. The task was to "select" the target, as quickly as possible, by positioning the display cursor over the circular target, and to confirm the selection with a left mouse button click. The cursor could either be visible or invisible, depending upon the control strategy being used (this is discussed later on), but in either case the target turned red as soon as the cursor was detected to be within the target

boundaries. The target boundary was the visible target area. Confirmation was only valid if the mouse click was received while the target was selected (highlighted). If the cursor moved out of the target area before confirmation, the target returned to white.

A trial ended upon successful target confirmation. When the target selection was successfully confirmed, the target flashed green for 100 msec, then disappeared, followed by immediate appearance of the next target. The start time of the new trial was the appearance of the new target. The process was repeated for the number of trials in the set.

Target sizes could be varied, but remained constant for a complete trial set.

search task

During the search task, all 20 targets were always displayed, as shown in figure 2. They appeared as white circles, which turned red whenever the cursor was within their boundary. A small number was displayed just above each target, starting with "1" for the upper left target, and progressing, sequentially, across each row, up to "20" for the lower right target. A small command box was displayed at the top center of the screen.

A trial set was initiated by a left mouse button click, which caused a number to appear in the command box. The task was to select the target whose number appeared in the command box, and confirm the selection with a left mouse button click. When the cursor moved within the boundary of any target, it turned red, whether it was the commanded target or not. Confirmation was only accepted if the left mouse button was clicked while the commanded target was selected (highlighted).

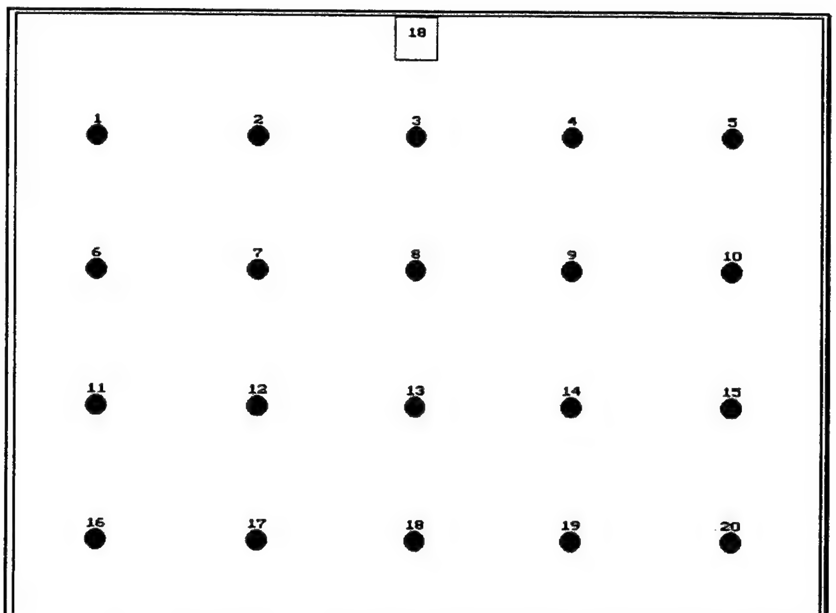


Figure 2. Search task display

A successful confirmation caused the target to flash green for 100 msec, followed by the appearance of a new command number in the command box. Start time for the new trial was the appearance of the new command. This sequence was repeated for the number of trials in the set.

The command box measured 0.5 inches square, on the 13 inch screen, subtending 1degree visual angle, at a 28 inch eye to screen distance. Numbers were 0.13 inch (0.25 degrees) high, both in the command box, and above each target.

There was no requirement for either eye fixation or cursor position to return to a starting point between trials. Of course the subject had to look to the command box to see the new command.

Target sizes could be varied, but were all the same size for a given trial set, and were always centered at the same 20 positions.

Cursor Control Types and Subjective Tests

Several different eye movement cursor control algorithms were tested subjectively, by the principle investigator and one colleague, using the tasks described in the previous section. The most straight forward algorithm, simply requires the subject to look at a display target to highlight the target, and to then press a confirmation button to select the target. The moving cursor is never visible to the subject. It is implemented by measuring point of gaze on the display, and placing the invisible cursor at that measured point. This will be referred to as "pure point of gaze control", or just "point of gaze control". As discussed previously, this can be successful only if point of gaze can be detected to an accuracy and precision smaller than the target width.

Attempts were made to enable selection of targets that were smaller than our measurement accuracy by using several different variations and enhancements of this basic strategy. The three major strategy categories were the following:

1. Make the cursor visible.
2. Make the cursor visible, and give the subject means to offset the measurement for correction of local errors.
3. Allow the subject to switch to a different type of control (not point of gaze control) for fine positioning, once the cursor is close to the target.

All strategies implemented were tested subjectively, by the principle investigator and one colleague. The control techniques and subjective results are discussed in the following subsections. A subset of these control techniques were selected for formal experiment trials and these are discussed further on.

Pure point of gaze control with and without a visible cursor

Point of gaze on the display was measured by a head mounted eye tracker integrated with a magnetic head tracking system. The system integrates head and eye position information to compute point of gaze on the display, and the cursor is positioned at the computed point of gaze. When target width is at least twice the accuracy of the system, this technique works extremely well. Subjective performance is actually best when the cursor is not displayed. The

user simply sees the target light up whenever he looks at it. When the cursor is displayed, under these conditions, it is an unnecessary distraction, and often increases the amount of time it takes to select a target.

So long as the eye tracker is working optimally, it seems to make little difference whether a fixation filter is used or not; but if there is occasional noise on the eye tracker measurement, the fixation filter can often prevent this from being noticed. Fixation filtering may slightly decrease the minimum target width for effective use of the open loop strategy, but this cannot be confirmed by the subjective testing that has been done so far.

When target size is significantly smaller than eye tracker accuracy, the pure open loop technique simply does not work. With the cursor continually displayed, the user can purposely look off target to correct errors, but this is an extremely unnatural and tiring task. When the target width approaches the precision (amount of jitter) of the eye tracker, the task becomes virtually impossible. Fixation filtering, averaging, and other types of low pass filtering seem to help only slightly.

Point of gaze control with visible cursor and freeze feature

The first feed back correction tested was the "cursor freeze" technique. When the right mouse button is held down an offset is continually computed to keep the cursor frozen in place. When the button is released, this offset is maintained and the cursor will continue from its "frozen" position. The user corrects local errors by freezing the cursor, looking directly at it, and then releasing it.

The freeze technique makes it practical to select targets that are about half the size of those required for dependable, pure point of gaze operation; however selection times feel substantially longer than standard mouse selection. The technique requires the cursor to be continually displayed, and takes some practice. The natural error is to move gaze away from the cursor (back to the target) just before releasing the button. This has the effect of negating any offset correction and creates the illusion that the offset correction is not working.

Control using eye-with-respect-to-head measure

A radically different technique, also tested, is to slave the cursor position not to point of gaze on the display, but rather to *eye position with respect to the head* (as measured by a head mounted eye tracker). The cursor is simply moved on the monitor as a function of eye position with respect to the head (the quantity actually measured by the head mounted eye tracker). Achieving a given cursor position requires adjustment of both head and eye position. In other words the user fixates a desired target, and while maintaining fixation on the target, makes head position adjustments to place the cursor on the target.

Once cursor position and point of gaze match, and so long as head position does not change, the cursor should follow point of gaze to within eye tracker accuracy limits. Head position becomes a continuous means of offset correction. Note that if a target, on the display, is fixated, and the head is

rotated towards the left, the cursor will move to the right. This is because the eyes have moved to the right *with respect to the head.*

It is easy, and quite natural for someone to make head movements while fixating a stationary point. We do this every time we read a sign while walking down the street. Of course, controlling a display cursor with such movements is by no means a natural or familiar activity.

There were high hopes for this technique because of its inherent simplicity, but it proved to be disappointing. Even with fixation filtering and other forms of low pass filtering, the cursor is too sensitive to small head/eye motions. It is excruciatingly difficult for the user to hold his head still enough or to make the tiny head movements needed to correct cursor position.

Usable cursor control was achieved with this method only by significantly reducing the gain between eye position and cursor movement. This means that if the user is looking at the cursor and changes his point of gaze (with head held stationary), the cursor will only move about 1/4 the distance that gaze is shifted. With this lower gain, it is possible to make fine enough head position adjustments to position and stabilize the cursor over small targets. However, because of the low gain, large head movements are needed to move the cursor long distances. If the target is at the edge of the screen, for example, the user winds up looking at the target out of the corner of his eye in order to achieve the desired cursor position.

Eye-with-respect-to-head plus freeze feature

The same sort of freeze feature previously described (cursor freezes while mouse button held down), can be added to the "eye with respect to head" method. It is implemented by continually computing an offset to balance any change in eye position, while the "freeze button" (right mouse button) is depressed. Upon release of the freeze button, the cursor continues form its current position.

The freeze feature allows positioning of the cursor without requiring unreasonable head positions. The user moves the cursor part way, freezes the cursor while moving his head back to a comfortable position, then moves the cursor the rest of the way.

Using this technique, the principle investigator and a colleague were both able to position the cursor over targets subtending about 1/4 degree visual angle. It remains, however, a very unnatural procedure that is annoyingly slow.

Pure point of gaze control plus user activated switch to head control

Very good subjective results were finally obtained with the following technique. The cursor position is computed with the pure point of gaze technique until the right mouse button is depressed. When the right mouse button is held down, the cursor moves proportionately to subsequent head rotation angles. It works best when the cursor is not displayed to the user until the mouse button is pressed. In effect, the user first searches for and fixates the target, then, if the target is not already highlighted, presses the right mouse button. When the

button is pressed, the cursor appears very near the target, and the user moves his head slightly to bring it on target.

Remember that the eye tracker is accurate to about a degree visual angle, so the cursor always appears within, or just barely beyond, the foveal area. Because the needed correction is always small, the gain for head movement control can be set fairly low, thus making it easy to achieve stable control, but never requiring uncomfortably large head motions. The subjective sensation is that a stable cursor can always be made to appear just off the target, and then moved on target with a very moderate head motion. After trying several values, head control gain was set to 0.5. In other words, the head must move twice as far as the desired cursor motion. This gain value is consistent with optimal head control gains found by Lin, Radwin, and Vanderheiden (1992).

If the cursor is displayed before depressing the mouse button, during the open loop control phase, its presence is annoying without being at all useful, and appears to degrade performance. The user quickly adopts the strategy of trying to ignore the cursor until the button is pressed.

This technique (with cursor invisible until the button press) allowed the principle investigator and a colleague to easily select the smallest targets included in the test program (0.25 degree visual angle diameter), with what seemed to be equal or greater speed than standard mouse selection. Subsequent quantitative experiments proved this to be illusory. As will be shown later, quantitative results show some trails that are as fast as the fastest mouse trials, but, on average, task completion time is slower than with standard mouse control.

Head position control

Head control was tested as a point of comparison with other studies. The cursor must be continually displayed for this method to work (as with a standard mouse). A couple of different gains were used. One gain produced a true head pointing result (the cursor appeared where an extension of the user's nose would intersect the screen), and the other gain was about half the first. Gain becomes a very important issue. If gain is too high, it is hard to position the cursor with enough precision. If gain is very low, an uncomfortably large head motion is required to move the cursor long distances on the screen. As previously mentioned, a gain of 0.5 is consistent with optimum head control gains found by Lin, Radwin, and Vanderheiden (1992). It should be noted that, as implemented for the current work, head control does not simply move the cursor proportionately to head azimuth or elevation, but actually computes the screen intersection of an imaginary vector attached to the subject's head.

The informal result was that pure head control did not seem as effective as the combined eye head technique previously described, but is a clearly viable technique for selecting targets of all the sizes tried.

Mouse control

Mouse control was used as a standard, against which to compare other control techniques, as well as for comparison to other studies. The mouse control parameters were set for "best feel" by an experienced mouse user. This turned

out to be a basic gain of about 4. An acceleration factor, as implemented by the standard Microsoft mouse driver, is included.

General comments

Several of the eye control techniques, described above, require the cursor to remain continuously visible to the user. A visible cursor slaved to point of gaze can be very disconcerting. We are used to being able to move our gaze with respect to objects in our visual field. When an object is slaved to point of gaze, and especially when it "dances" about a point slightly displaced from point of gaze, it is extremely distracting. It also creates an enormous urge to look towards the cursor, causing the cursor to move farther away (positive feedback), sometimes resulting a game of "chase the cursor".

A fixation filter can be added to any of the techniques involving eye control. Fixations are the periods during which the gaze point is relatively stable, and during which most visual information is received. During typical scanning behavior, fixations are connected by very rapid eye jumps, called saccades. Very little, if any visual information is acquired during saccades. Eye position data can be processed to try to exclude miniature eye movements, and measurement noise, during periods of fixation, and allow the measurement to change only in response to saccades. The potential advantage is a quieter more stable cursor. The potential disadvantage is some additional lag between control input and cursor response, and coarser control resolution.

A simple, on line fixation algorithm allows fixation position to change only when a specified number of sequential point of gaze samples have less than a specified standard deviation. The position of the new fixation is set to be the mean of those points. Generally, the number of sequential points is set to represent about 100 msec, usually considered to be the minimum time needed to acquire visual information. The standard deviation value is set to be larger than the standard deviation of miniature eye movements that occur during fixations, and also larger than expected measurement noise during fixations. A standard deviation value of 0.5 degrees is typical.

Fixation filters were tested with all eye movement control techniques. Fixation filtering did not appear to bring dramatic improvement to any of the techniques, although it may have produced small improvements, not obvious during subjective testing. When eye position data is subject to brief periods of very large noise, usually due to intermittent feature recognition failure by the eye tracker, fixation filtering can dramatically improve results (see Jacob, 1990). The equipment and environment used for the current study resulted in extremely stable recognition of the pupil and corneal reflection (the features used by the eye tracker to compute eye position), and violent noise due to recognition failure never appeared.

Fixation filtering was not included in formal experiment trials.

Experiment design

Formal experiment trails were performed using 4 different cursor control strategies:

1. standard mouse control ("mouse")
2. pure point of gaze control ("pog")
3. point of gaze with a user controlled switch to head control for fine positioning ("pog&hd")
4. head position control ("head")

All of these control strategies were implemented as previously described.

Each subject performed 9 "Fitts" and 9 search task trial sets, using each control type. Each Fitts trial set consisted of 30 trials with the same target diameter, but a pseudo-random pattern of movement distances. Each target appeared at one of 20 pre-defined positions, arranged in an evenly spaced grid pattern of 5 horizontal by 4 vertical positions.

The nine trial sets consisted of three trial sets with each of 3 different target diameters. For all trial sets other than pog control sets, the 3 target diameters were 0.25 inches, 0.5 inches, and 1 inch. These sizes correspond to 0.5, 1, and 2 degrees visual angle, respectively, at the 28 inch eye to screen distance used. For pog control trial sets, the target diameters were 1, 1.5 and 2 inches (2, 3, and 4 degrees visual angle), during Fitts task trials; and 1, 1.2, and 1.4 inches (2, 2.4, and 2.8 degrees visual angle), during search task trials. The 1.5 and 2 inch targets could not be used for the search task, where all targets were displayed at once, because adjacent targets would overlap. Target sizes smaller than 1 inch were not used during pog control trials because the accuracy capabilities of the eye-head tracker would have been far exceeded.

Each search task trial set consisted of 20 trials. The target order was a random sequence of the 20 possible targets. As in the case of the Fitts task, the target size remained constant during each trial set.

Fitts trial sets and search trial sets were not intermixed. After all trial sets for one task type were completed, the subject was given a break, and then trial sets were performed with the alternate task. Each subject performed a total of 270 Fitts trials and 180 search trials for each type of cursor control.

Data was recorded for all trial sets, although it was intended that the first 3 trial sets for any given control type be considered practice.

Figure 3 shows the distribution of possible distances between targets, defined by the 20 target positions (400 possible origin and destination pairs). The pseudo random target sequence for the Fitts task was constrained by a requirement that there be an equal number of trials in each of 5 distance ranges. The ranges, expressed in pixels, were <200, 200-300, 300-400, 400-500, and >500.

On the 13 inch display used for the experiment, there were approximately 67 pixels per inch. It was discovered, late in the data collection process, that most VGA monitors, including the one being used, have a significant scaling non-linearity. On the display used, the number of pixels/inch varied by about 10% from one screen edge to the other. The implications for experiment results are discussed later on.

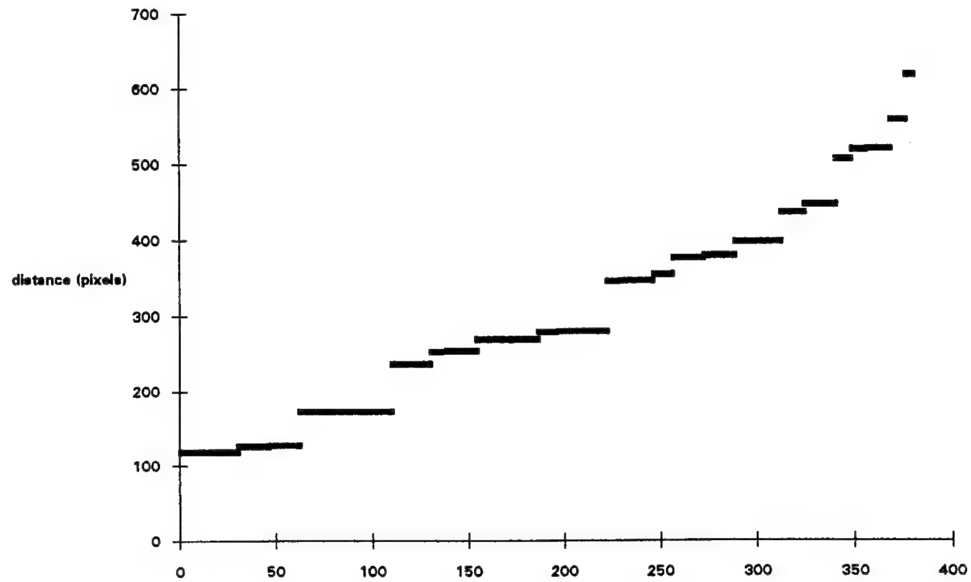


Figure 3. Distances between all possible target pairs, for Fitts task

Data was collected from 6 subjects for the Fitts task, and 5 subjects for the search task. (Search task data from one of the 6 subjects was lost due to a computer recording error).

Every trial set was recorded as an individual ASCII data file on the display computer. Data recorded for each trial included start time, starting cursor position, target number and position, previous target number and position, time and cursor position when the target was first selected, and time and cursor position when the target was confirmed. In addition, the time and cursor position for every mouse button press or release, and a "cursor fixation" list were also recorded for each trial. The cursor fixations were the result of a simple, on line fixation algorithm than operated on computed cursor position. It was not intended to describe eye fixation behavior, but rather to provide an abbreviated description of cursor movements.

Point of gaze data, as computed by the eye-head tracker system were recorded, separately, on the eye tracker computer. Each trial set was also recorded as a separate file. Flags were sent to the eye tracker computer by the head tracker computer when each trial started, and when each trial ended. These flags were recorded along with point of gaze data. Point of gaze data consists of eye point of gaze coordinates, pupil diameter, and distance from the eye to the point of gaze, sampled every 1/60 of a second. Note that this is the same data that was sent, in real time, to the display computer. During head control trial sets, data actually consisted of the head vector intersection with the display, rather than eye point of regard.

The signal from the fixed scene camera was video taped for all runs. The video tapes show the display monitor viewed by the subject, with a superimposed

cursor showing point of gaze computed by the eye-head tracker system. During head control trials, the superimposed cursor indicates intersection of a head fixed vector with the display surface.

RESULTS

Based on accuracy measurements for the eye tracker, it was anticipated that targets larger than 1 degree visual angle radius could be easily identified by the open loop point of gaze measure. In other words, when a person fixated a target larger than 1 degree radius, under the point of gaze cursor control paradigm, it should have immediately turned solid red. As data was being collected, it was noticed that this was not always the case. Sometimes a target highlighting flickered on and off, and subjects found that they had to make a slight head position or gaze adjustment to get a stable red target highlight. This is the reason for some obvious outliers in the pog control data.

The unexpected cursor placement errors are partially explained by the display screen non-linearity, mentioned in the previous section. The eye-head tracker system measures point of gaze with respect to an absolute coordinate system on the display screen surface. The display computer converted these coordinates to pixel positions with a simple gain and offset computation. Failure to account for the display screen non-linearity could easily have resulted in 0.25 inch errors in cursor placement on some parts of the screen. This effectively made point of gaze measurements less accurate by up to 0.5 degrees visual angle.

Another factor that may have diminished the effective open loop accuracy of the eye-head tracker was the uniformity of the large targets. No visible feature was provided to mark the center of these target disks. To the extent that subjects may not have fixated the center of the target, measurement system errors displaced the cursor from a position that may already have been off center.

It was noticed from observations during data collection, and subsequent observations of video tapes, that subjects take an unexpectedly long time to switch from pog to head control, when using the pog&hd technique. The eye-head tracker point of gaze computation is visible to an observer (not the subject) as a set of cross hairs on the eye-head tracker system "scene monitor". Looking at this monitor, point of gaze can be seen to move to the target, and remain there for longer than seems reasonable before the right mouse button is activated to visualize the cursor on the subject's display. It is not a hardware delay affect. there is no more than 17 msec between activation of the mouse button and appearance of the head controlled cursor.

A related observation concerns a common mistake made when using the pog&hd control technique. Subjects sometimes press the head control switch before finishing their saccade to the target, or at least before the eye tracker has reported the end of the saccade. Remember that there is up to a 50-67 msec delay between an eye movement event and availability of that information to the display computer. If a subject does press the switch too soon, the cursor appears at a position corresponding the beginning or mid point of the eye saccade, far away from the target. The subject must then either make a very

large head motion to bring the cursor on target, or release the head control switch and depress it again, while fixating the target. An attempt to guard against this mistake may sometimes cause the switching delay described in the previous paragraph.

Fitts task results

Since there was no requirement to return the cursor to a standard position between trials, index of difficulty values were calculated based on the distance between actual cursor position at the start of a trial and the target position. Note that this is often different from the distance between the previous target (nominal cursor origin) and the current target. Index of difficulty (ID) is therefore more of a random variable, rather than a fixed set of values. ID is still closely correlated with the nominal previous-to-current target distance.

Plots of sequential trial completion times, normalized by Fitts index of difficulty, show little evidence of learning effects, beyond the first trial set (first 30 trials), for any of the control techniques.

Figure 4 shows target distance plotted versus task completion time for the point of gaze control data. The regression line slope (4.71 msec/degree) is not significantly different from zero. The intercept, pulled up by all of the outliers, is 696 msec, but the minimum values are in the 200-250 msec range.

Figures 5 through 8 show scatter plots of Fitts' index of difficulty versus task completion time for the 4 different control techniques. Data from all subjects are

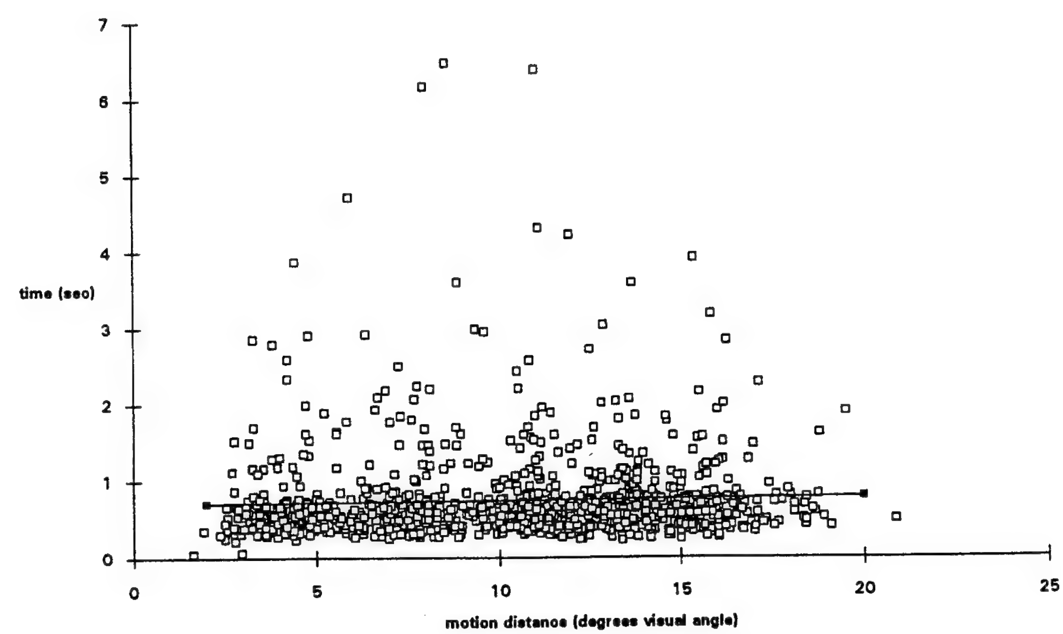


Figure 4. Task completion time versus motion distance for point of gaze control data

included, but only the last 6 trial sets, for each control type, are included for each subject. Even though learning effects were not apparent beyond the first trial set, the first 3 trail sets (one set at each target size), are considered practice. Each scatter plot, therefore, contains 1080 data points (30 trials/set, times 6 trial sets, times 6 subjects).

Several things are immediately obvious. There is a great deal of variance in all of the data, but more for the pog and pog&hd control cases than for mouse or head control. There are also some obvious outliers in the pog and pog&hd data. The data tend to curve up at low ID values, as often observed in Fitts' law data (MacKenzie, 1992). Variance seems to increase with ID value (heteroscedasticity). In the case of pog control, the minimum values (fastest completion times) do not seem to increase with increasing ID. The mean, or median, on the other hand, does increase.

If we use Welford's ID, as shown in figures A1 through A4, in Appendix A, the data show less tendency to curve up at low ID values. A similar affect can be observed with the Shannon formulation, although these plots are not included. As shown in figures A5 through A8, in Appendix A, the heteroscedasticity disappears if the predicted variable is log time, instead of time.

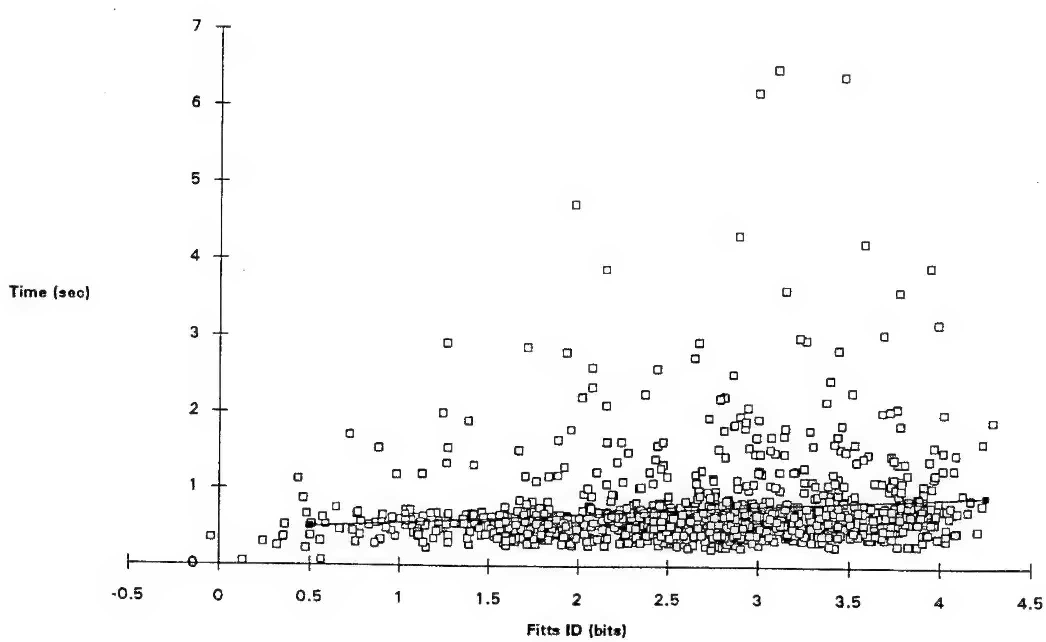


Figure 5. Point of gaze control, trial completion times plotted versus Fitts index of difficulty

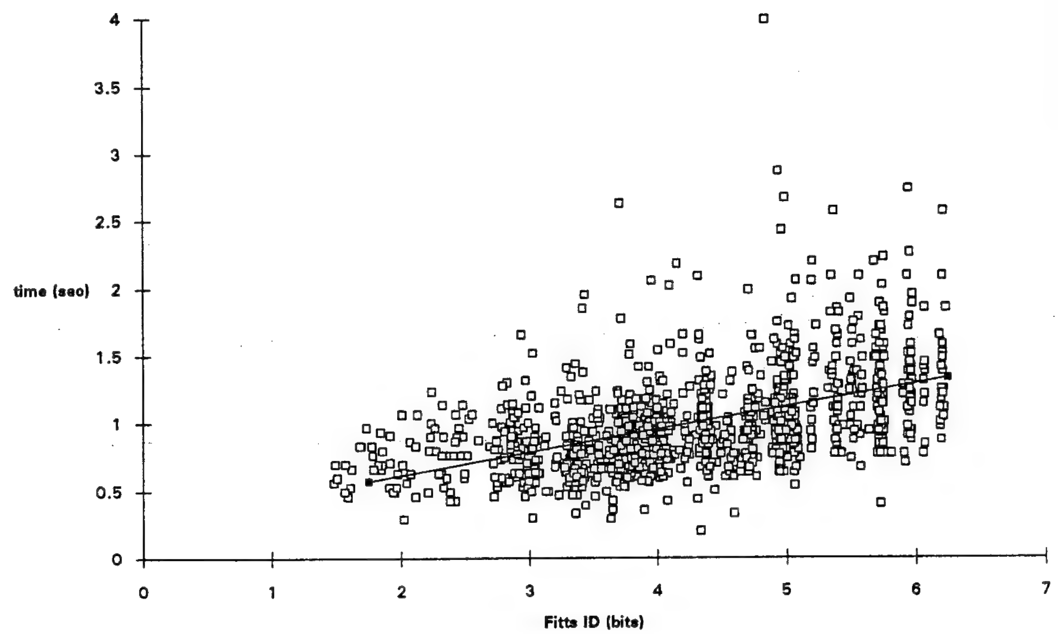


Figure 6. Mouse control, trial completion times versus Fitts index of difficulty

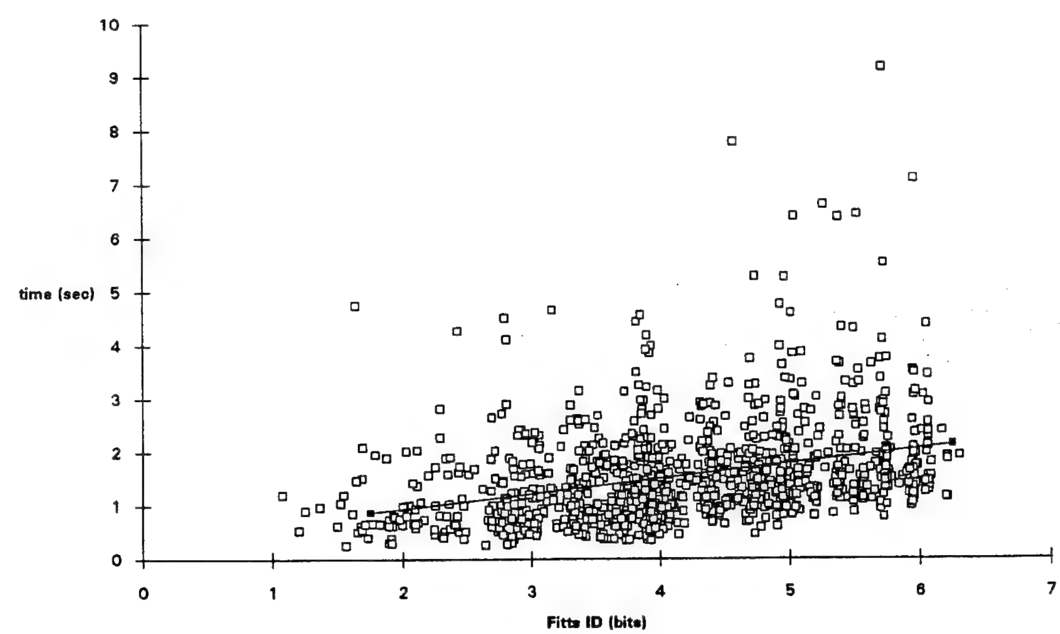


Figure 7. Point of gaze with switch to head control, trial completion times versus Fitts ID

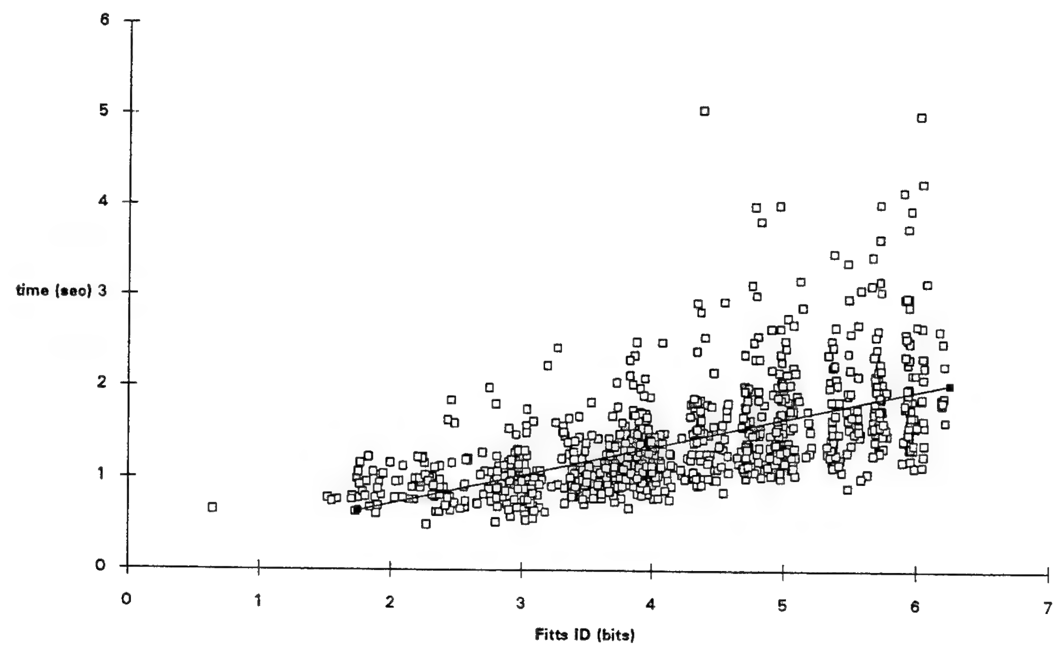


Figure 8. Head control, trial completion times versus Fitts index of difficulty

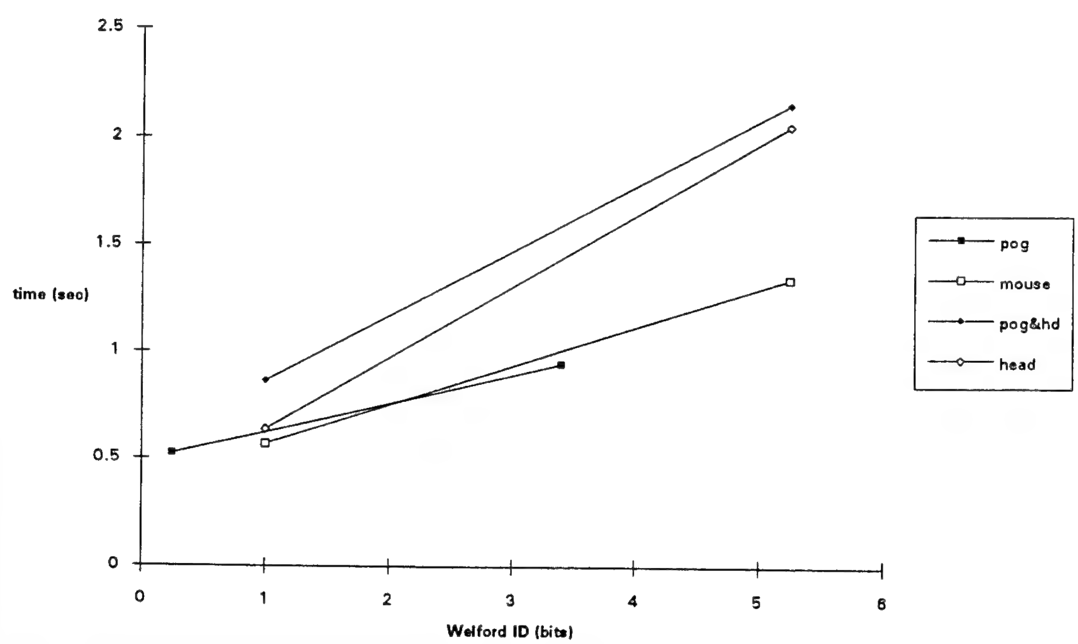


Figure 9. Regression lines for all control types, based on raw data, and Welford ID

Figure 9 shows regression plots corresponding to the scatter plots discussed above. Regression was performed on raw data as shown in the scatter plots, not mean or median data. An F test shows the regressions all to be significant at $\alpha < .001$, but correlation coefficients are very low.

Tests for parallelism and for equal adjusted means (equality of intercept values) show that head and pog&head control regressions do not have significantly different slopes, but do have significantly different intercepts, but only at the $\alpha = .1$ level. The pog and mouse control regression lines are not significantly different from each other. The pog and mouse regressions <u>are</u> significantly different from the head and pog&hd regression lines ($\alpha < .001$).

Although the mean pog&hd task completion times, as indicated by the regression line, are slightly greater than those for head control, minimum pog&hd times are similar to the fastest mouse times. Head control, does not show any completion times approaching the fastest mouse times. This is illustrated by figure 10. The figure shows minimums for all of the control types calculated by partitioning the data into ID bins, each 0.5 bits wide, and then selecting the minimum from each bin containing at least 25 samples. This may imply that faster responses are possible (although not usual) with pog&hd than with head alone. On the other hand, the fast pog&hd responses may simply be cases for which the eye tracker put the cursor on target, and target selection did not require activation of the head control switch. This could be confirmed by a more detailed analysis of the data.

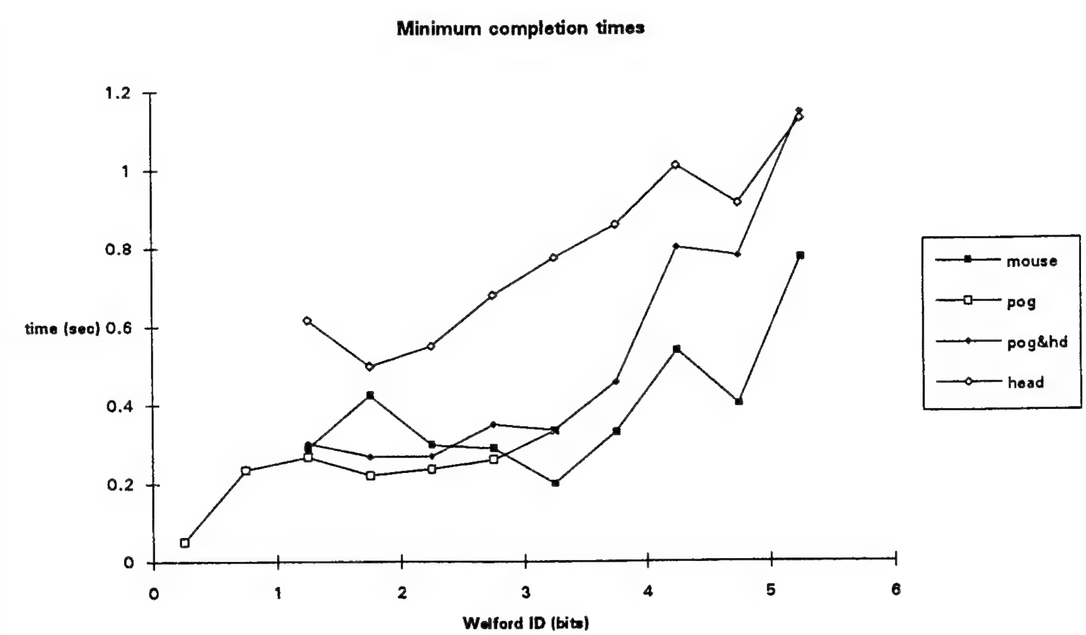


Figure 10. Minimum trial completion times, within 0.5 bit ID ranges

In most Fitts law literature, regressions are reported not on raw data, but on mean or median values computed for each ID level. Sometimes individual means or medians are computed for each subject, and sometimes data from all subjects are pooled. In other cases it is difficult to determine which of these techniques were used. While this should show valid relationships between the variables, information on the variance of raw data is lost. Regressions on mean or median data can be expected to have much higher correlation coefficients than regressions on raw data.

To enable more equivalent comparison with other published data, data from the current study were also analyzed in terms of median values. Median, as opposed to mean values were selected in order to better eliminate the affect of outliers.

Because the current study did not have a fixed set of ID values, computation of median values is not as straight forward as it other wise would be. Data were grouped with respect to ID ranges. Medians were then computed for samples falling in each range (or bin). The bin widths were set at 0.5 bits, starting from ID = 0. In other words, bins were ID = 0-0.5, 0.5-1.0, 1.0-1.5, etc. The median for each bin was plotted at an ID value equal to the mid range of the bin. Data from all subjects were pooled, and as with the raw data, the first 3 trial sets for each control type were excluded. Median values were used only when computed from more than 5 samples; in other words, when there were more than 5 samples in the bin.

The results of the median value analysis are shown in figure 11. Regression lines are qualitatively similar to those for the raw data, but slopes and intercepts

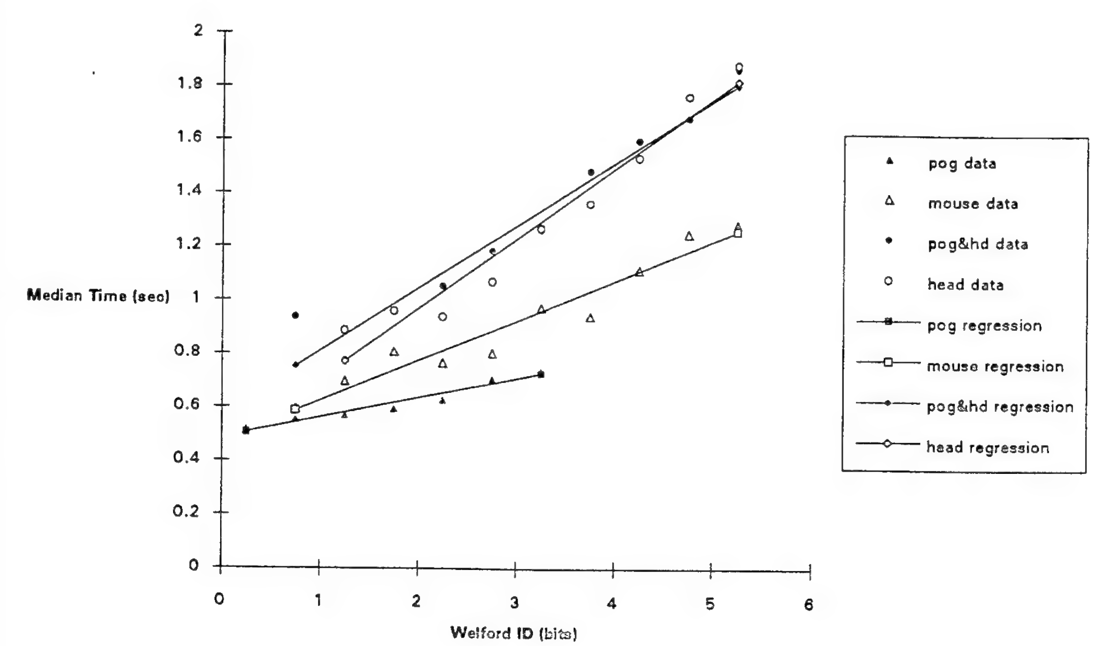


Figure 11. Median trial completion time data and regression lines, using the Welford ID

are smaller, and correlation coefficients are very high, as would be expected. The regression lines are good fits to the "dense" sections of the Figure A1-A4 (Appendix A) scatter plots, ignoring the outlying points. Table 1 shows slopes, intercepts, correlation coefficients (r), and coefficients of determination (r^2), for the regression data in the current study, and from other studies in the literature.

MODEL: motion time = intercept + slope * ID; IP = 1/slope							
Study	**control**	**intercept (s)**	**slope (s/bit)**	**IP(bits/s)**	**r**	**r^2**	**comments**
current	mouse	0.276	0.169	5.90	0.48	0.23	*Fitts ID, raw data*
current	mouse	0.388	0.182	5.49	0.48	0.24	*Welford ID, raw data*
current	mouse	0.394	0.136	7.34	0.95	0.91	*Fitts ID, median data*
current	mouse	0.474	0.149	6.71	0.97	0.94	*Welford ID, median data*
Epps,1986	mouse	0.108	0.392	2.60	0.83	0.69	*Welford ID*
Card, et. al., 1978	mouse	1.03	0.096	10.40	0.91	0.83	*Welford ID*
Radwin, 1990	mouse	-0.06	0.147	6.80	0.99	0.98	*Fitts ID*
Lin, et.al., 1992	mouse	0.125	0.135	7.40	0.99	0.99	*Fitts ID*
current	pog	0.45	0.11	9.02	0.16	0.02	*Fitts ID, raw data*
current	pog	0.49	0.134	7.47	0.16	0.03	*Welfor ID, raw data*
current	pog	0.42	0.076	13.19	0.93	0.87	*Fitts ID,median data*
current	pog	0.487	0.073	13.71	0.97	0.95	*Welford ID, median data*
Ware & Mikaelian, 1987	pog	0.68	0.073	13.70	NA	NA	*Welford ID*
current	pog&hd	0.379	0.282	3.54	0.33	0.11	*Fitts ID, raw data*
current	pog&hd	0.562	0.304	3.29	0.33	0.11	*Welford ID, raw data*
current	pog&hd	0.306	0.247	4.05	0.99	0.98	*Fitts ID, median data*
current	pog&hd	0.576	0.234	4.27	0.97	0.95	*Welford ID, median data*
current	head	0.109	0.309	3.24	0.60	0.36	*Fitts ID, raw data*
current	head	0.306	0.333	3.00	0.60	0.36	*Welford ID, raw data*
current	head	0.312	0.24	4.17	0.97	0.95	*Fitts ID, median data*
current	head	0.44	0.263	3.81	0.97	0.95	*Welford ID, median data*
Radwin, 1990	head	-0.096	0.24	4.18	0.97	0.95	*Fitts ID*
Lin, et.al., 1992	head	0.25	0.169	5.92	0.97	0.94	*Fitts ID*
Jagacinski & Monk, 1985	head	-0.268	0.199	5.00	0.99	0.98	*Fitts ID*

Table 1. Regression model parameters for current study, and from other studies in the literature. Values for the Ware and Mikaelian (1987) study are for the "button press confirmation" protocol, and are the values derived by MacKenzie (1992) from plots in the Ware and Mikailian paper. Values for the Radwin (1990) study were computed by averaging values reported for different motion directions.

Search task

The search task was a more realistic computer interaction task, involving a limited search component, but does not easily lend itself to as detailed a quantitative analysis. It would be difficult, for example to compute an index of

difficulty, for each trial, that would allow fair comparison of control types. In the case of mouse control, for example, the user had to look to the command box between trials, but did not have to move the cursor to the command box. This makes the motion distance parameter somewhat ambiguous. It could be taken as distance from the previous target to the new target, or distance from cursor position at the time the subject fixates the command box, or distance from the command box to the new target. For point of gaze control, the situation is rather different, since the cursor does follow the scan path. It is also difficult to factor in difficulty of the actual search component for individual trials, since the location of some points may be better remembered than others.

It is reasonable, however to assume equivalent difficulties for entire trial sets, or groups of trial sets, since each trail set included all target positions and a single target size. Figure 12 shows median trial completion times, for all subjects, as a function of target size. To present the data in a form that is more consistent with the index of difficulty formulation the figure 12 plot shows median completion times versus $log_2(1/target\_size)$.

As with the Fitts law task, and contrary to expectations, the pog&hd technique was not only much slower than mouse control, but was also slightly slower than pure head control. A more detailed analysis, breaking completion times into component parts, would be instructive and is discussed in the next section.

It is also noteworthy that, unlike the Fitts task results, point of gaze control was slower than mouse control, even at similar ID levels. This may be due to a couple of factors. As previously discussed, delay was sometimes introduced by

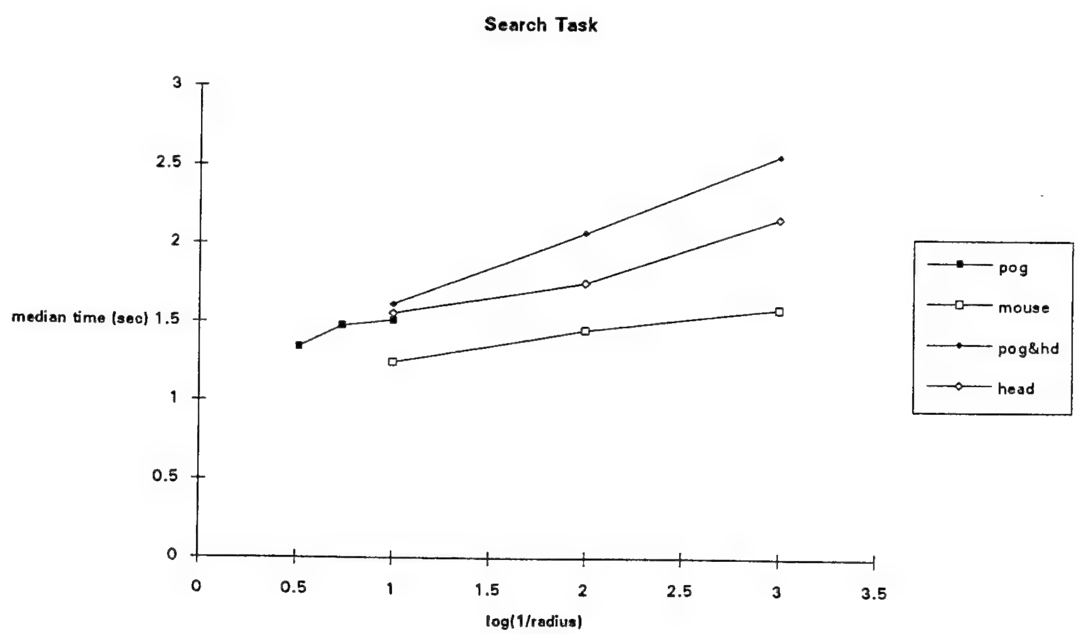


Figure 12. Median search trial completion times versus base 2 log of inverse target radius

errors in point of gaze computation that approached the size of the targets. This may have happened more often during the search task, for unkown reasons. There may also have been mouse motion, tending to decrease cursor to target distance, during the search phase of the mouse control task. Analysis of the separate point of gaze data, in conjuntion with the task performance data, might provide a more definitive explanation.

DISCUSSION

Mouse control is clearly the fastest and easiest of the methods tested for selecting display targets, if we consider the entire range of target sizes used. Point of gaze control is even faster and easier than the mouse, when the point of gaze measurement is accurate enough and dependable enough. With technology that is currently available and practical, for this application, that implies relatively large targets. This limitation will become less severe as the technology for point of gaze measurement improves, but may never reach the resolution possible with a mouse or other mechanical controller.

Implementation of the point of gaze technique, as used in the current study, could be improved by properly accounting for display non-linearity's, and by providing a visible point to mark the center of display targets.

Pure point of gaze control is best implemented without a visible cursor. It is better to highlight display elements when point of gaze is determined to fall on them (or sufficiently near to them). As suggested by Jacob (1990), selection can be a fixation duration criteria for non critical selections, or a separate confirmation switch for more critical selections.

Point of gaze control with a switch to another fine tuning modality is clearly viable when targets must be small, and when a mouse or similar mechanical controller cannot be used. As implemented in this study, however, it was not shown to be faster than pure head control, even for long motion distances. Point of gaze control, with a switch to head control, does have the advantage (over pure head control) of not requiring extreme head positions, since the head control is only used for fine tuning. The same basic technique could also be implemented with a switch to some other modality beside head control, for the fine positioning function. The disadvantage to the technique is that it carries the cognitive burden of switching modalities. It is a more complicated task than simple head or mouse control.

It should be expected that, when required motion distances are very short, the pog&hd technique would show longer task completion times than pure head control. In the case of a target distance that is the same as expected eye tracker error, for example, the head control phase of the pog&hd control should be the same task as pure head control, but implemented after an extra mode switching delay. As distance to the target is increased, we might expect to reach a point where time saved, by the quick cursor movement to within eye tracker error distance (point of gaze control phase), is greater than the mode switching delay. In other words, when plotting time against index of difficulty, we might expect to see larger intercepts but smaller slopes, for pog&hd data compared to pure head control. The current experiment trials either never

reached large enough motion distances to overcome the switching delay, or this model is incorrect. The data do tend to show larger intercepts, and slightly smaller slopes for pog&hd, than for pure head control, with a convergence point near the maximum IDs used, but these differences are not statistically significant.

It is possible that the technique of using point of gaze, plus another fine tuning mode, could be significantly improved over the results documented in the current study. The problem of switching modes too soon (causing the cursor to appear far from the target) might be reduced if a fixation algorithm is used to lock out switching during saccades. It may also be helpful to program the system for a short delay, after the head control switch is activated, before switching to head control. Such a programmed delay is likely to be shorter than conscious delay by a user. Any reduction in system transport delay would be beneficial as well. A preliminary step is additional analysis of the data already gathered, to look at relative time spent on the different phases of the task. This may shed some light on whether mode switching delays really are a significant factor, as postulated above.

As previously discussed, eye saccades are known to have a duration that is almost linearly related to amplitude. The amplitude range required for the Fitts task is about 4-18 degrees visual angle. Corresponding saccade duration's would be predicted to be about 35-76 msec. The targets used for the point of gaze control trials are large enough that we would not expect multiple saccades to be required to reach a target. We would expect trial confirmation times to have a 35-76 msec component that is a function of distance to the target, and much longer delays that are not closely correlated with target distance. These include the 50-100 msec period required to initiate a saccade in response to the appearance of a target, 50 msec transport delay in the eye tracker, and on the order of 100 msec to press the confirmation button once the target is highlighted. Thus, we might expect minimum task completion times to be 200 - 250 msec. Variance in the components not related to target distance could easily swamp the saccadic variation that is a function of target distance, and this is what appears to happen (see figure 4). Variance does increase with ID, probably accounting for positive linear regression slopes in figures nn, and nn. It must be noted, however, that the log(time) formulation (figure nn) does not exhibit such obvious increasing variance, yet has a statistically significant slope.

All of the data gathered in the current study, even the mouse data, show what seems to be rather large variance. Since other Fitts law studies examined in the literature do not report variance of the raw data, it is not clear whether or not this is typical. The serial nature of the task, and the transmission delays in the system may contribute to the variance, and to the tendency of variance to increase with index of difficulty. As often reported, use of the Welford or Shannon formulation, for index of difficulty, produces a better regression line fit than the original Fitts index. A still better fit (higher correlation coefficient, and more constant variance) is achieved by using *log(time)* as the predicted variable. This is consistent with findings by Sheridan and Ferrell (1963), who tested a remote controller with large transmission delays.

The pog and pog&hd control data show more variance than mouse or head control. Noise in the point of gaze measurements, and variance in mode

switching behavior for the pog&hd control data, probably account for the difference.

As noted by MacKenzie (1992), there is wide variation in Fitts law model parameters, between different studies reported in the literature. Differences are presumably due to differences in the precise nature of the tasks, differences in control devices, and a host of other unknown and uncontrolled differences. In the current study, for example, the mouse control task used higher gain than other studies, and used an acceleration term, unlike many other studies. Furthermore, a serial task was used, in contrast to the discrete tasks often used, and the distribution of motion distances was different. None the less, results presented in table 1 fit well within the range of other data reported for mouse and head control.

Looking at the raw (as opposed to median) data for point of gaze control, significant regressions were obtained for the Fitts law model, but the correlation of task completion time to Fitts ID is very weak (correlation coefficient = 0.16). The same can be said for the Welford, Shannon, and *log(time)* formulations. Much of the correlation that does exist might be explained by inaccuracy in the point of gaze control system. For smaller targets, there was a larger probability of the cursor being intermittently (as opposed to solidly) within the target boundary, resulting in longer task completion times. It was originally intended that targets, for the point of gaze control task, be significantly larger than measurement system accuracy limits. As previously described, unexpected sources of error added to those limits. It would be very interesting to see whether a cleaner implementation of this task would show any significant Fitts law relation. It should be noted, however, that Fitts law parameters for the point of gaze control data <u>are</u> reasonably consistent with the one other study available for comparison (Ware and Mikaelian, 1987).

More data was gathered in the current study than has yet been thoroughly analyzed. Relatively little use has so far been made of the mouse button activity data, time values for first target selection (as opposed to confirmation), or the separately recorded point of gaze data. The mouse button activity data and additional time data can be used to analyze task completion by phase, as previously discussed for pog&hd control. The separately recorded point of gaze data can be used to look at scan patterns independently of cursor control. This might help determine the cause for the apparent mode switching delays observed during pog&hd control trials. The point of gaze data can also be used to examine the relation of scan pattern to cursor position during mouse control, and to separate search times from other phases of the search task trials.

REFERENCES

Abrams, R. A. (1989). Speed and Accuracy of Saccadic Eye Movements: Characteristics of Impulse Variability in the Oculomotor System, *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 15, No. 3, 529-543.

Borah, J. (1989). *Helmet Mounted Eye Tracking for Virtual Panoramic Display Systems - Volume I: Review of Current Eye Movement Measurement Technology.* AAMRL-TR-89-019, Harry G. Armstrong Aerospace medical Research laboratory, Human systems Division, Air force Systems Command, Wright-Patterson AFB.

Card, S. K., English, W. K., & Burr, B. J. (1978). Evaluation of mouse, rate-controlled isometric joystick, step keys, and text keys for text selection on a CRT. *Ergonomics*, 21, 601-613.

Epps, B. W. (1986). Comparison of six cursor control devices based on Fitts' law models *Proceedings of the 30th Annual Meeting of the Human Factors Society*, 327-331. Santa Monica, CA: Human Factors Society.

Fitts, P.M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, 47, 381-391.

Jacob, R. K. (1990). What You Look At is What You Get:: Eye Movement Based Interaction Techniques, *CHI90 Proceedings*, 11-18.

Jagacinski, R. J., & Monk, D. L. (1985). Fitts' law in two dimensions with hand and head movements. *Journal of Motor Behavior*, 17, 77-95.

Lin, M. L., Radwin, R. G., & Vanderheiden, G. C. (1992). Gain effects on performance using a head-controlled computer input device. *Ergonomics*, 35(2), 159-175.

MacKenzie, I.S. (1992). Fitts' Law as a Research and Design Tool in Human-Computer Interaction. *Human Computer Interaction*, 7, 91-139.

Radwin, R. G. (1990). A Method for Evaluating Head -Controlled Computer Input Devices using Fitts' Law. *Human Factors*, 32(4), 423-438.

Sheridan, T. B., & Ferrill, W. R. (1963). Remote manipulator control with transmission delay. *IEEE Transactions on Human Factors in Electronics*, 4, 25-29.

Ware, C., & Mikailian, H. H. (1987). An evaluation of an eye tracker as a device for computer input. *Proceedings of the CHI+GI '87 Conference on Human Factors in Computing Systems and Graphics Interface*, 183-188.

Welford, A. T. (1960). The measurement of sensory-motor performance: Survey an reappraisal of twelve years' progress. *Ergonomics*, 3, 189-230.

Young, L. R., & Sheena, D. (1988). Eye Movement Measurement Techniques. in Encyclopedia of Medical Devices and Instrumentation, ed. Webster, J., John Wiley & Sons, Inc. New York, 1259-1269.

APPENDIX A. SCATTER PLOTS

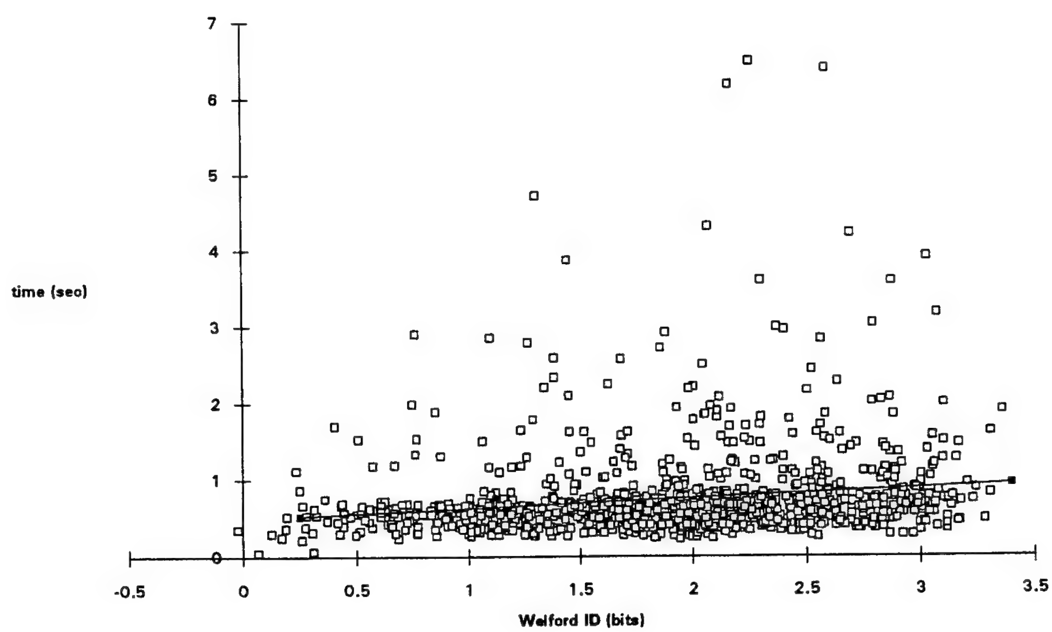


Figure A1. Point of gaze control, trial completion times versus Welford index of difficulty

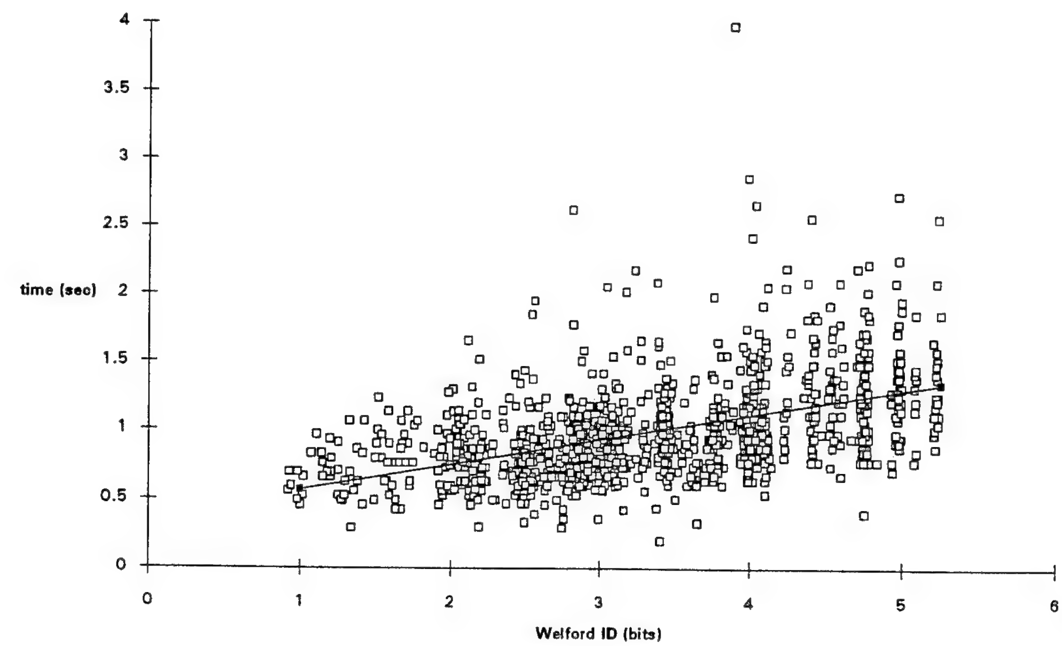


Figure A2. Mouse control, trial completion times versus Welford index of difficulty

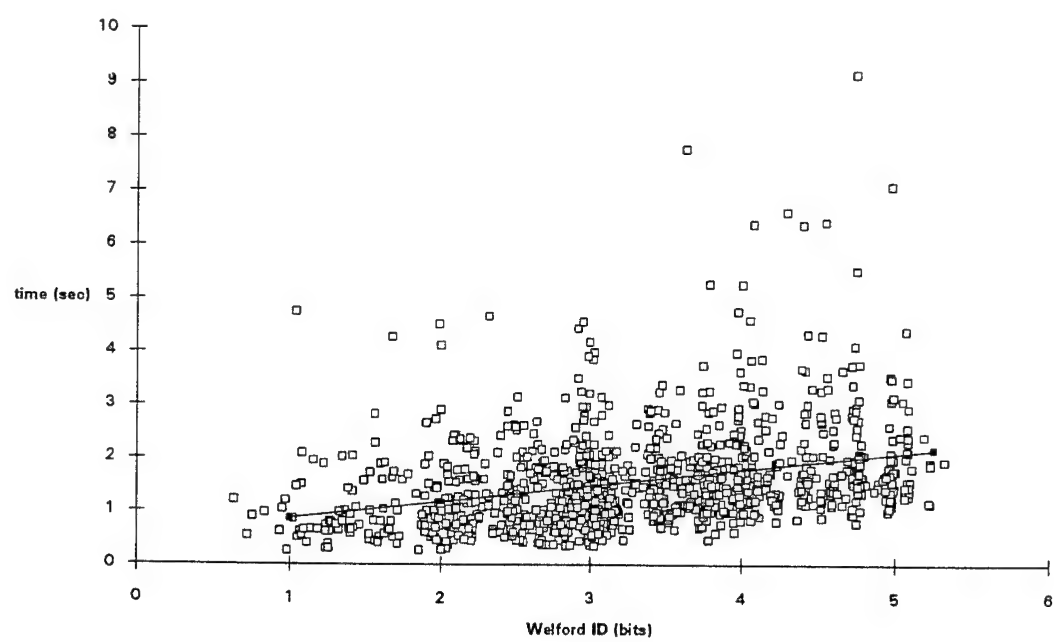


Figure A3. Point of gaze with switch to head control, trial completion times versus Welford index of difficulty

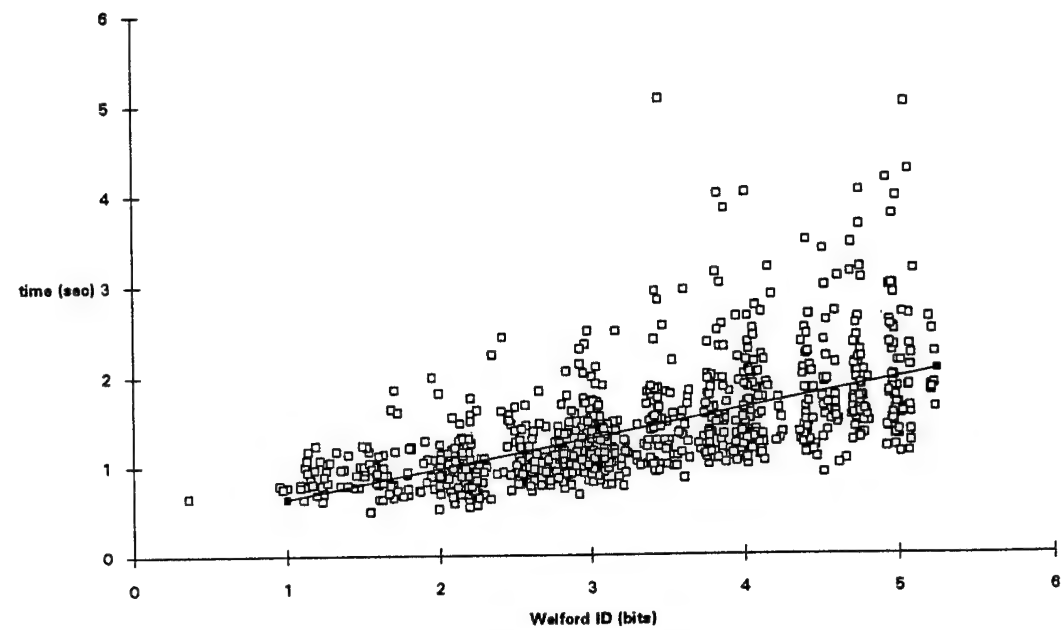


Figure A4. Head control, trial completion times versus Welford index of difficulty

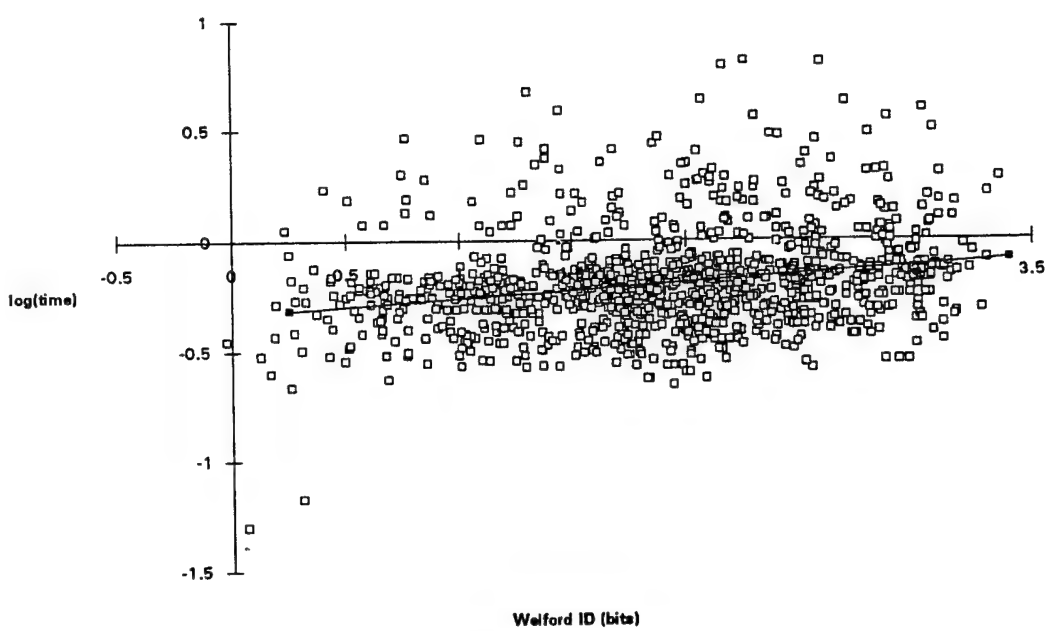


Figure A5. Point of gaze control, log of trial completion time versus Welford ID

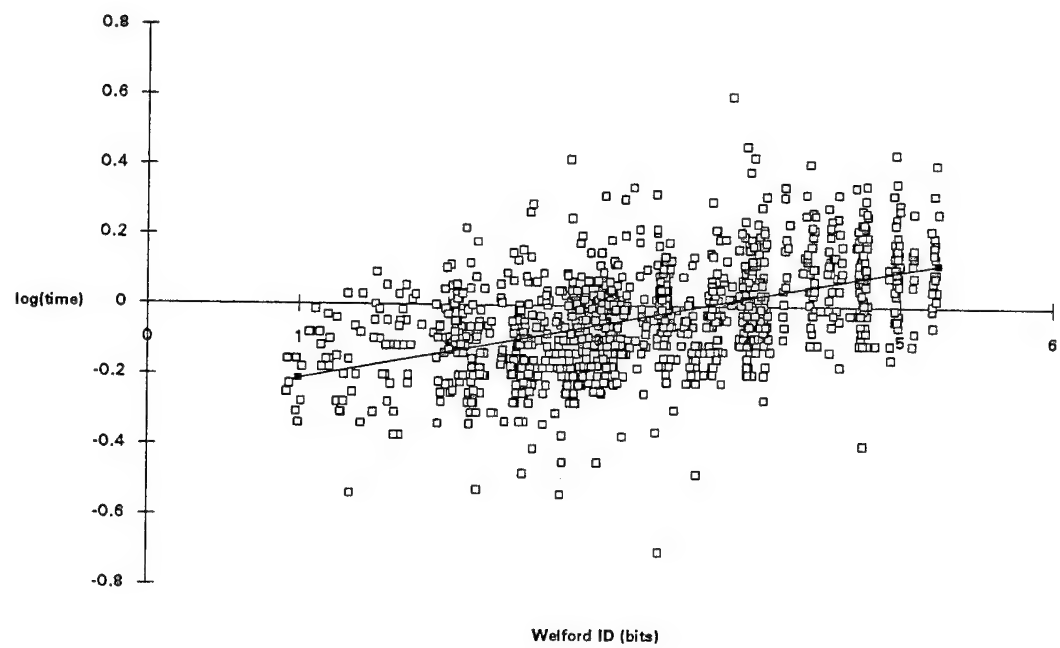


Figure A6. Mouse control, log of trial completion time versus Welford ID

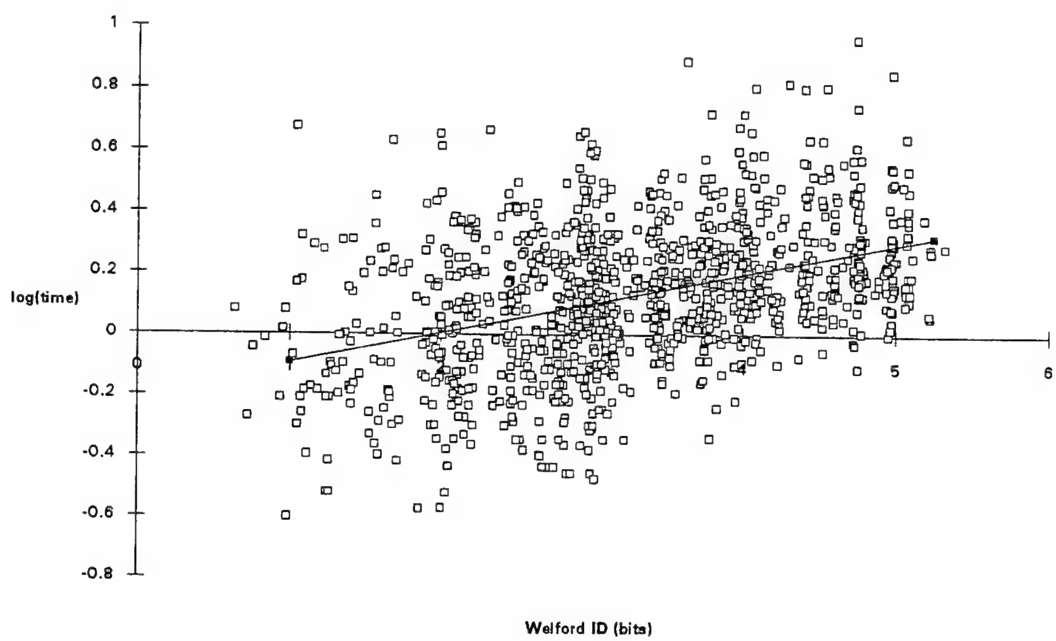


Figure A7. Point of gaze with switch to head control, log of trial completion times versus Welford ID

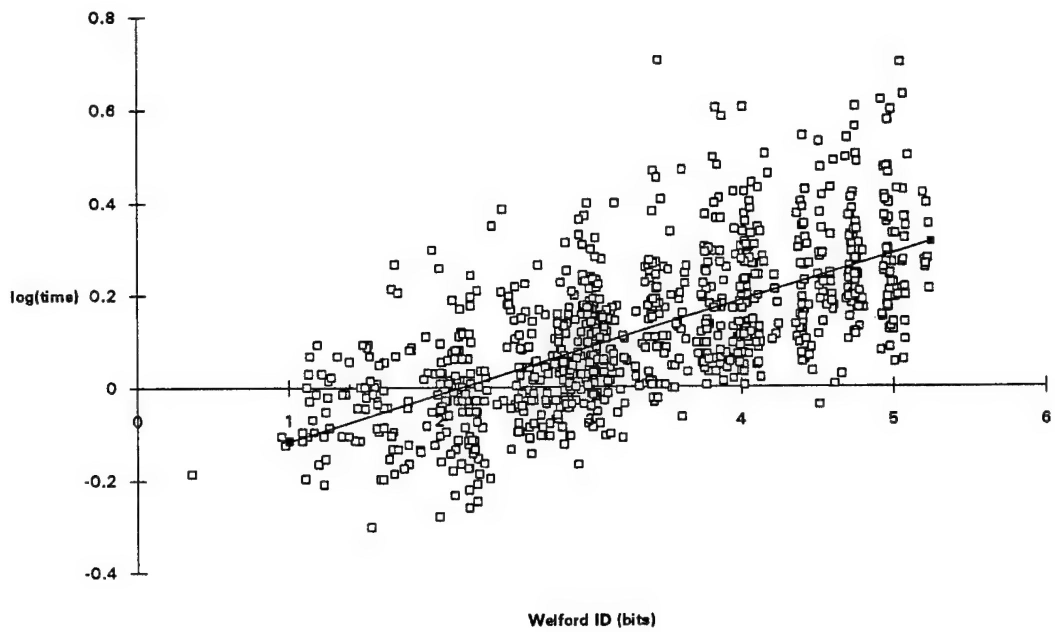


Figure A8. Head control, log of trial completion time versus Welford ID

Investigation of Eye and Head Controlled cursor Positioning Techniques

Addendum to Final Report
June 1995
contract no F41624-93-C-6006

Joshua Borah
Applied Science Laboratories
Bedford, MA

The *pog&hd* cursor control technique proved slower than anticipated. To better understand this result, the Fitts' task, *pog&hd* control data have been examined in further detail than that presented in the final report.

Postulated reasons for longer than expected completion times include:

- mode switch activation too early (before completion of saccade to target), resulting in need for multiple tries, or extra long head control phase;
- cognitive and manual switching time (perhaps influenced by attempt to avoid mistake described above).

The raw data includes a mouse button history for each trial, and fixation list for each trial. A program was written to extract the following values from these lists:

- The number of times the mode switch (right mouse button) was activated before target confirmation;
- The time, from the beginning of the trail, that the mode switch was last activated before target confirmation;
- The position of the cursor at final mode switch activation;
- Distance of the cursor, at the time of final mode switch activation, from the target center;
- Index of difficulty (Fitts, Welford, and Shannon formulations) based on distance from cursor position at mode switch activation to target center.
- The first fixation (time and position), prior to mode switch activation, that was within 0.5 inch (1 degree visual angle) of the cursor position at mode switch activation, with no intervening fixations outside the same 0.5 inch boundary.

The last of these items defines the time when point of gaze arrived at the position it was to maintain, with out significant change, until mode switch activation. The time from this fixation to mode switch activation was used as an estimate of mode switching delay. It is an imperfect estimate because of the somewhat arbitrary selection of the 1 degree visusal angle boundary, and because it does not account for the delay period between mode switch activation and the actual beginning of head control activity.

All of the regression data discussed below is based on raw data, rather than median data for various ID ranges. The full affect of variance is preserved, and the results seem more accurately descriptive of the data. Although the other alternatives were computed, it is the Welford ID formulation that is presented below. The data, as presented below, is all based on the last 6 trails for each subject and each control technique.

There were 1080 “Fitts task” trials (counting only the last 6 trials for each subject) using the *pog&hd* control technique. In 19% of these trials, the mode switch was never used. The subject was able to select the target (make it turn red) with point of gaze alone, and pressed the confirmation button without ever changing to head control mode. These trials are concentrated at the lower ID values because this most often occurred for the larger target sizes.

In 10% of the trials the mode switch was pressed more than once. These were usually cases in which the switch was activated too soon (before the system registered fixation on the target) and the cursor appeared too far from the target. The subject released the switch (returning to point of gaze mode), fixated the target, and pressed the switch again.

In some cases, even though the cursor appeared very far from the target when the head control switch was activated, the subject used head control to finish the task rather than releasing and re-activating the switch. If we arbitrarily define "too far" as greater than 1 inch (2 degrees visual angle), then about 7% of the trials fall into this category.

The mean cursor to target distance when the mode switch was pressed (the final time, in cases where it was activated more than once) was 0.62 inches (1.24 degrees visual angle). The standard deviation was 0.64 inches (1.28 degrees visual angle).

The mean and standard deviation for the mode switching delay, as defined above, is shown in figure 1. Note that subject 5 had a substantially longer mean delay than other subjects. The group mean was 0.49 seconds with a standard deviation of 0.52 seconds. If subject 5 is excluded, the group mean falls to 0.35 seconds with a standard deviation of 0.3 seconds.

In order to reduce the number of confounding variables we can look at the subset of pog&hd trials that <u>did</u> use the mode switch, but that <u>did not</u> require multiple mode switch activation or require head controlled positioning over more than 1 inch. In other words, we will look at the runs for which the two phase control technique was used in a consistent, nominally ideal fashion. The word "ideal" was used in data plot legends to describe this subset of *pog&hd* data.

The "ideal" subset turns out to be 65% of the original 1080 runs. As shown by figure 2, the resulting regression line for completion time versus Welford ID shows a smaller slope and higher intercept than the line for all *pog&hd* runs. This is to be expected. The runs that used only point of gaze tend to have relatively short completion times and low ID values, and omitting these runs tends to increase the average time at low ID values.

The point of gaze control phase of each trial was examined by plotting ID versus time, from the trial start to mode switch activation. The head control phase was examined by plotting ID versus time, from mode switch activation to trial completion. In the latter case, ID is computed using distance to the target center, from the cursor position at the time head control is activated.

Linear regression plots in figure 3 show nearly the same slope for the head control phase of the *pog&hd* data and the *head* control data. The approximately 200 msec difference in the intercepts may be due to a delay between mode switch activation and the actual beginning of head motion. If this is so, then 200 msec should be added to the delay observed before switch activation (figure 1), yielding a mean mode switching delay on the order of 0.5 sec. Statistical significance of the intercept difference has not yet been computed.

The negative ID values for some of the head phase data points show trials in which the target was already selected (red) when the mode switch was activated.

Always activating the mode switch, regardless of whether the target is already highlighted, is a more reasonable strategy than may be apparent. Although some time is wasted when point of gaze alone has succeeded, it eliminates the cognitive decision time required to notice whether the target is "lit".

The pog phase of the pog&hd data is quite similar, although not identical to the pog control trials. Statistical tests, to determine significance of the differences, have not yet been performed.

The scatter plot in figure 4 shows quite a few outliers even for the set of "ideal" trials. Trials with confirmation times greater than 4 seconds were examined individually. They all appear to be cases for which the subject simply spent a long time in an unsuccessful attempt to highlight the target without resorting to the head control switch. The subject eventually gave up the attempt to highlight the target by shifting point of gaze and activated the head control switch. All of these trials were from two of the subjects (4 and 5), and in all of these cases the target diameters were 0.5 or 0.25 inch, less than the smallest targets explicitly used for *pog* trials (1 inch diameter).

In these extreme cases, it is obvious that subjects held on to a poor strategy for too long; but in general the data provide no way to measure the use of this strategy. The outliers described above are too few to have a large affect on overall results.

It seems likely that the unexpectedly long average completion times for the *pog&hd* trials are largely attributable to an extremely long mode changing delay. If the delay between fixation and mode switch were removed, the "ideal" pog&hd data would be faster than head control for ID values above 2.2. If an additional 200 msec were removed (possible delay after switch activation, as previously discussed), then the "ideal" pog&hd data would be essentially the same as mouse control.

The reasons for such long mode switching delays can not easily be teased from the existing data. We can postulate that the delays have a significant cognitive component and may be exacerbated by attempts to avoid switching too early.

Smaller affects are attributable to variations in strategies used by particular subjects at particular times. Most of these secondary affects are eliminated by looking only at the "ideal" trials as described above. Some exceptionally short trial completion times can be attributed to successful attempts to use point of gaze alone. In these trials the eye tracker proved accurate enough to properly indicate the target when the target was first fixated. Some exceptionally long trails can be attributed to <u>unsuccessful</u> attempts to use point of gaze alone. In these trials the eye tracker did not indicate the target when fixated. Some other trials were unusually long because the subject activated the head control switch too soon

Future experiments can attempt to use intelligent switching logic to lock out the head control switch during saccades. Modifications to improve accuracy of the point of gaze component, as described in the final report, may also prove significant. Another possible variation on the *pog&hd* control technique would be

to switch modes automatically during long fixations, and automatically switch back when a long saccade is detected.

In order to generate data with less ambiguity, it might be interesting to simply instruct subjects to use the head control switch whether needed or not. This would eliminate trials that really use the *pog* technique, would eliminate any cognitive time spent noticing whether the switch was needed, and would prevent the poor strategy of attempting to use point of gaze alone for targets that are smaller than eye tracker accuracy.

The difference between the mode switching delay times for subject 5 and other subjects is not explained by anything in the recorded data. Although the data do not seem to show a continuing learning curve, it would, none the less, be interesting to see the affect of significantly longer practice. It is possible that a learning curve would be evident over a longer period. It certainly seems evident that strategy is important.

Finally, it would be interesting to gather subjective data to evaluate preference between a mode switching scheme and pure head motion control. Both seem potentially viable for situations in which manual control is not possible. The former carries a potentially annoying cognitive burden, while the latter requires large head motions that may prove annoying and tiring over time.

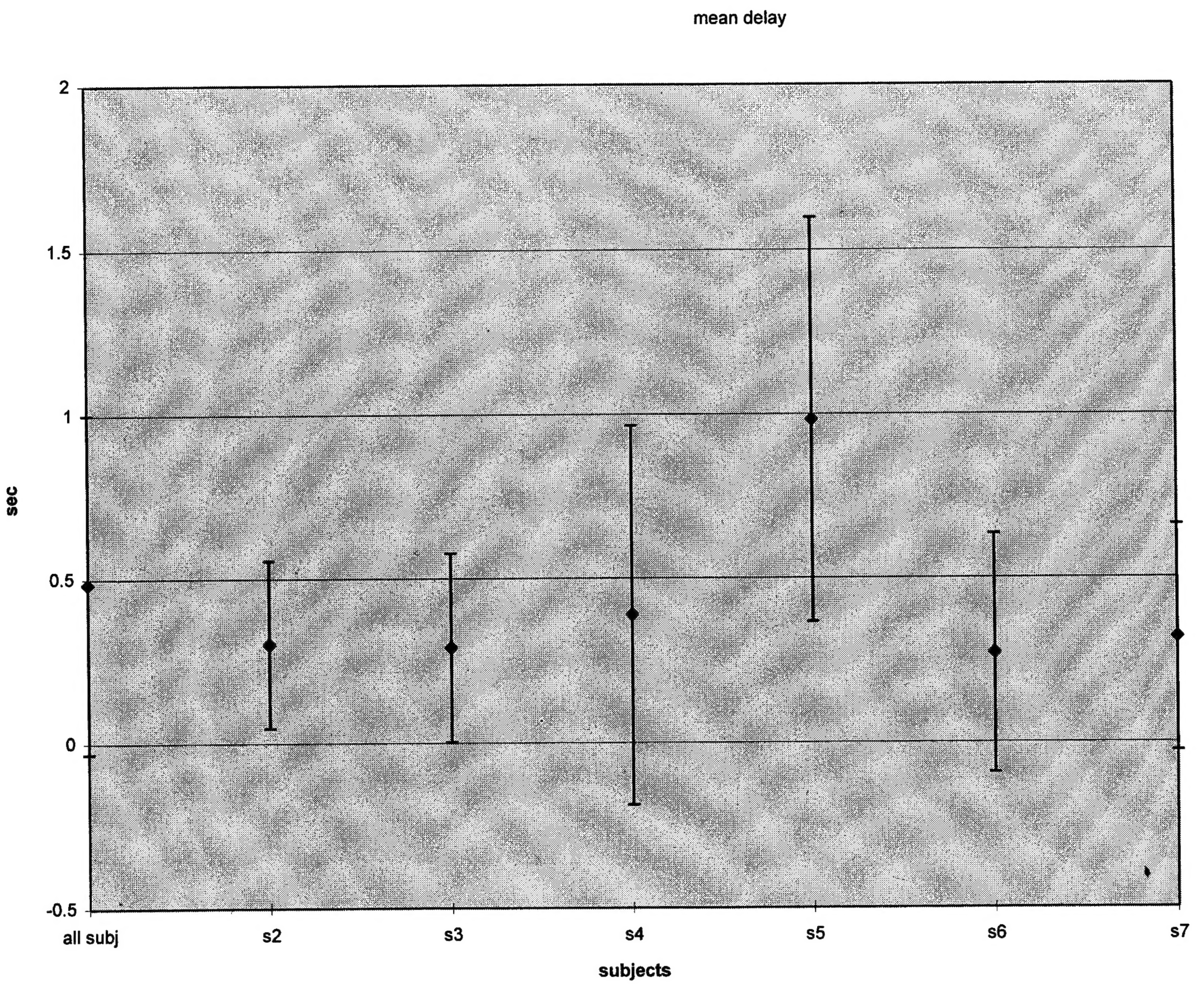


FIGURE !.

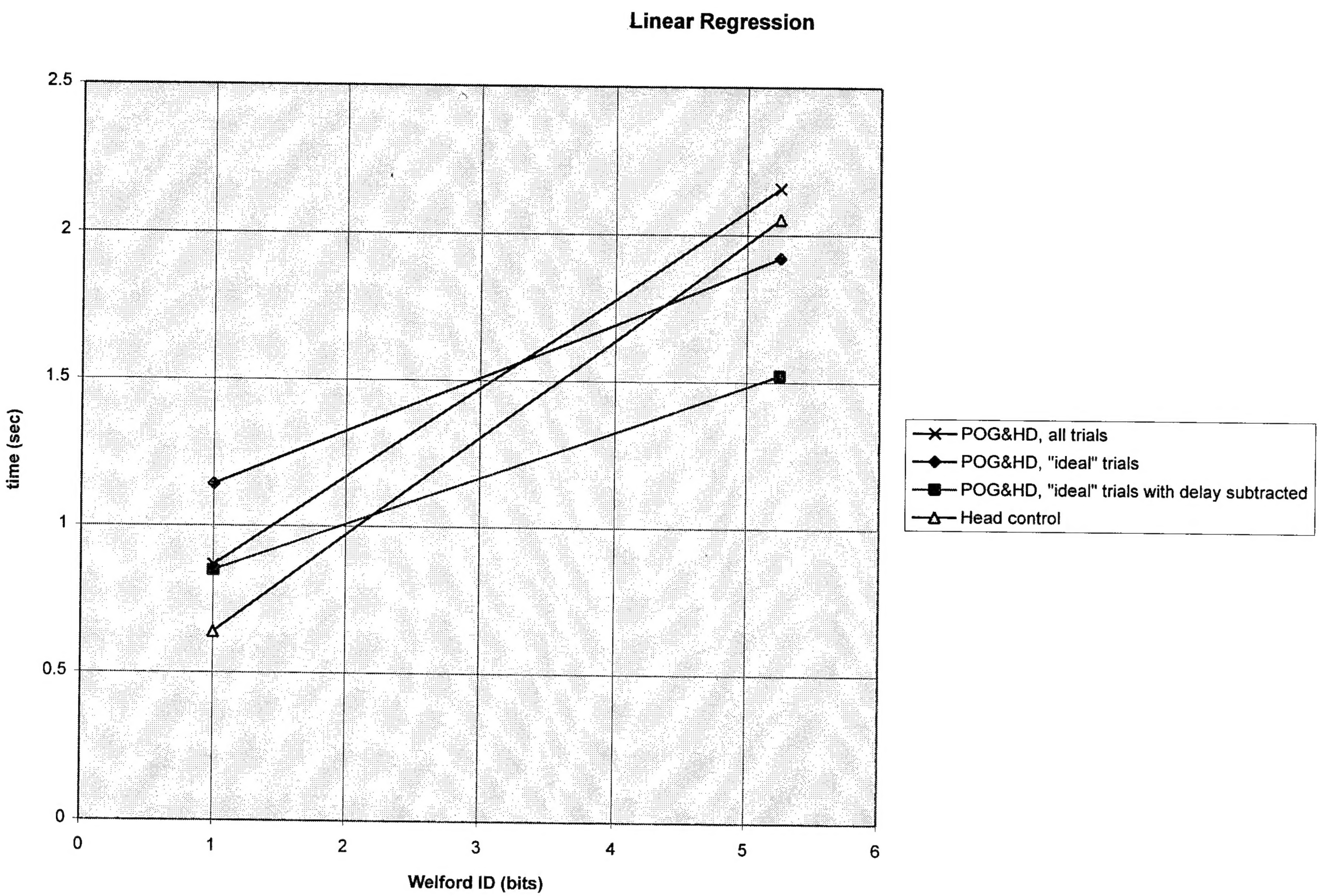

Linear Regression
time (sec)
2.5
2
1.5
1
0.5
0
0
1
2
3
4
5
6
Welford ID (bits)
POG&HD, all trials
POG&HD, "ideal" trials
POG&HD, "ideal" trials with delay subtracted
Head control

FIGURE 2.

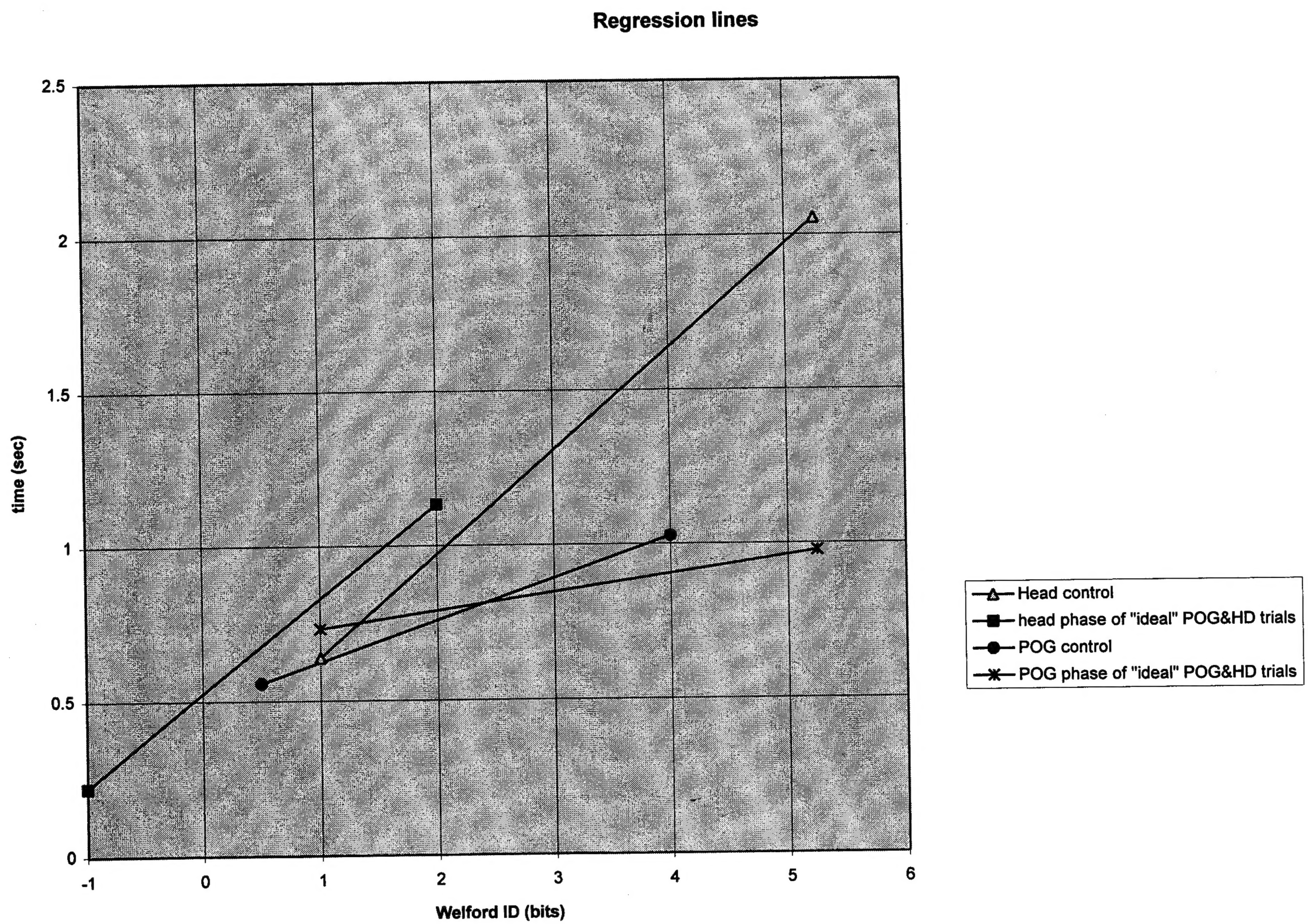
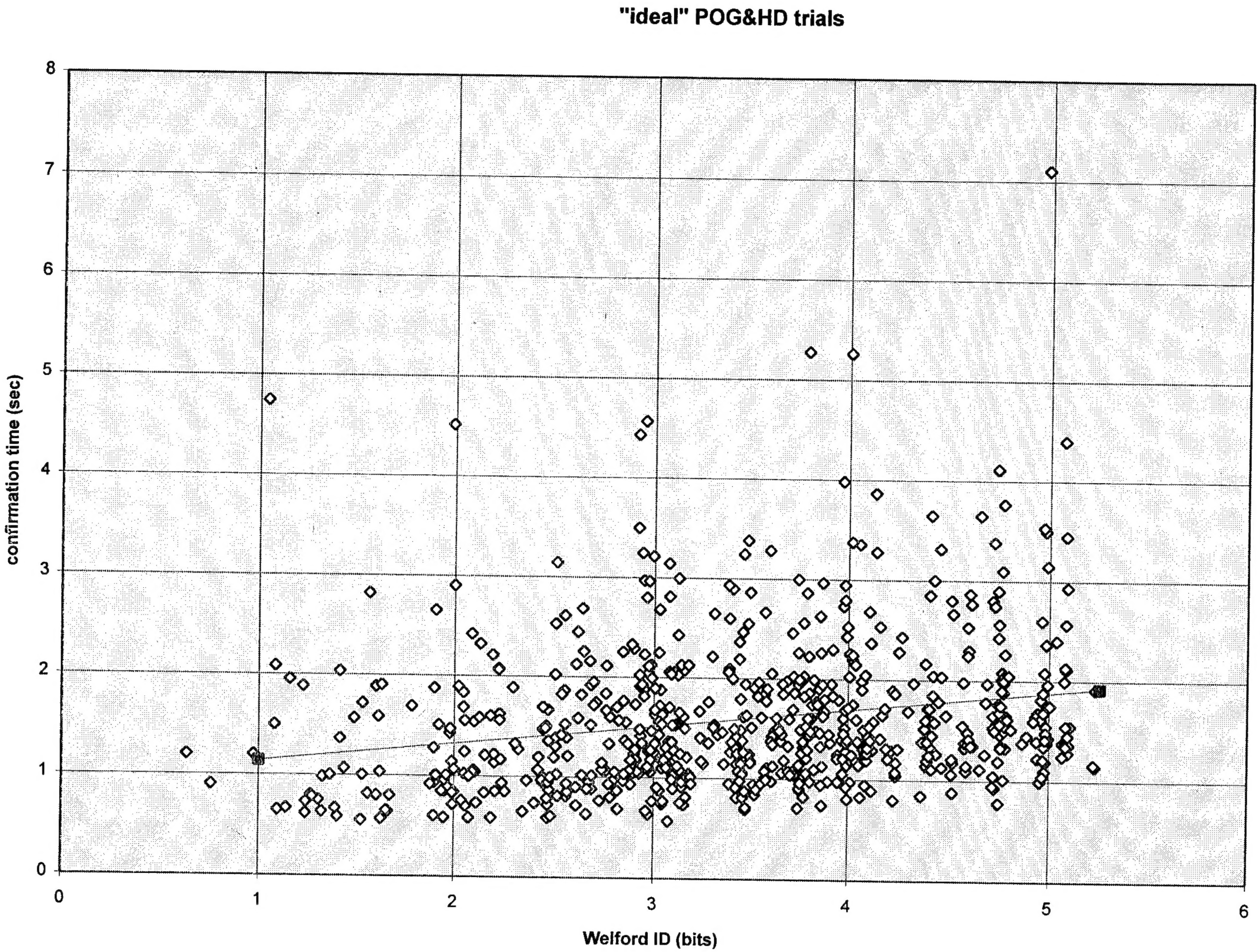

Regression lines
time (sec)
2.5
2
1.5
1
0.5
0
-1
0
1
2
3
4
5
6
Welford ID (bits)
Head control
head phase of "ideal" POG&HD trials
POG control
POG phase of "ideal" POG&HD trials

FIGURE 3.



U. S. Government Printing Office 1996 750-071/00168

FIGURE 4.